a topographic map. A stack of such sections, drawn on transparencies, yields a three-dimensional electron density map (Fig. 7-37*b*). Modern structural analysis, however, is often carried out with the aid of graphics computers, on which electron density maps are contoured in three dimensions (Fig. 7-37*c*).

Protein Crystal Structures Exhibit Less Than Atomic Resolution

The molecules in protein crystals, as in other crystalline substances, are arranged in regularly repeating three-dimensional lattices. Protein crystals, however, differ from those of most small organic and inorganic molecules in being highly hydrated; they are typically 40 to 60% water by volume. The aqueous solvent of crystallization is necessary for the structural integrity of the protein crystals as J. D. Bernal and Dorothy Crowfoot Hodgkin first noted in 1934 when they carried out the original X-ray studies of protein crystals. This is because water is required for the structural integrity of native proteins themselves (Section 7-4).

The large solvent content of protein crystals gives them a soft, jellylike consistency so that their molecules lack the rigid order characteristic of crystals of small molecules such as NaCl or glycine. The molecules in a protein crystal are typically disordered by a few angstroms so that the corresponding electron density map lacks information concerning structural details of smaller size. The crystal is therefore said to have a resolution limit of that size. Protein crystals typically have resolution limits in the range 2 to 3.5 Å, although a few are better ordered (have higher resolution, that is, a lesser resolution limit) and many are less ordered (have lower resolution).

Since an electron density map of a protein must be interpreted in terms of its atomic positions, the accuracy and even the feasibility of a crystal structure analysis depends on the crystal's resolution limit. Figure 7-38 indicates how the quality (degree of focus) of an electron density map varies with its resolution limit. At 6-Å resolution, the presence of a molecule the size of diketopiperazine is difficult to discern. At 2.0-Å resolution, its individual atoms cannot yet be distinguished, although its molecular shape has become reasonably evident. At 1.5-Å resolution, which roughly corresponds to a bond distance, individual atoms become partially resolved. At 1.1-Å resolution, atoms are clearly visible.

Most protein crystal structures are too poorly resolved for their electron density maps to reveal clearly the positions of individual atoms (e.g., Fig. 7-37). Nevertheless, the distinctive shape of the polypeptide backbone usually permits it to be traced, which, in turn, allows the positions and orientations of its side chains to be deduced (e.g., Fig. 7-37*c*). Yet, side chains of comparable size and shape, such as those of Leu, Ile, Thr, and Val, cannot be differentiated with a reasonable degree of confidence (hydrogen atoms, having but one electron, are not visible in protein X-ray structures), so that a protein structure cannot be elucidated from its electron density map alone. Rather, the primary structure of the protein must be known, thereby permitting the sequence of amino acid residues to be fitted, by eye, to its electron density map. Mathematical refinement can then reduce the errors in the crystal structure's atomic positions to around 0.1 Å (in contrast, the errors in small molecule X-ray structure determinations may be as little as 0.001 Å).

Most Crystalline Proteins Maintain Their Native Conformations

What is the relationship between the structure of a protein in a crystal and that in solution where most proteins normally function? Several lines of evidence indicate that *crystalline proteins assume very nearly the same structures that they have in solution:*

1. A protein molecule in a crystal is essentially in solution because it is bathed by solvent of crystallization over all of its surface except for the few, generally small patches that contact neighboring protein molecules. In fact, the 40 to 60% water content of typical protein crystals is similar to that of many cells (e.g., see Fig 1-13).

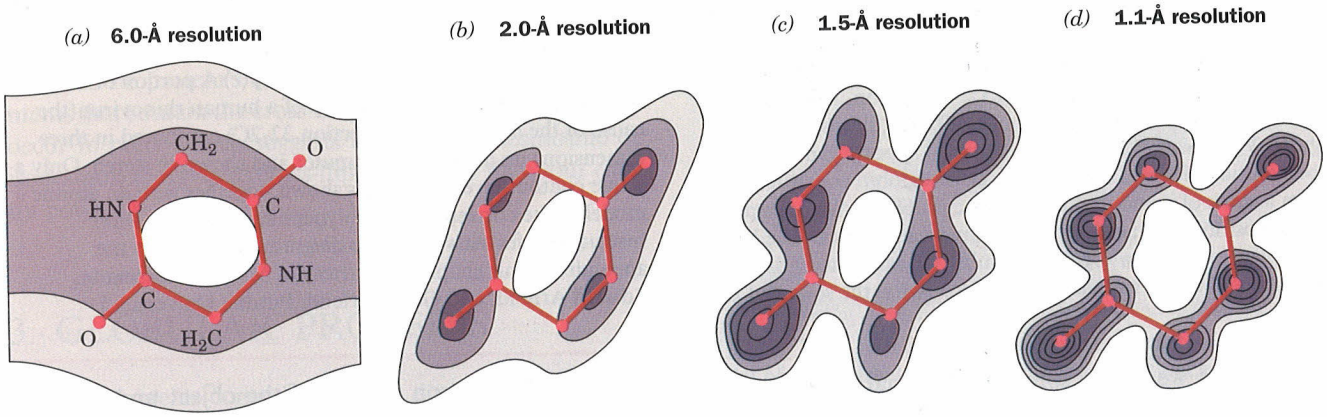
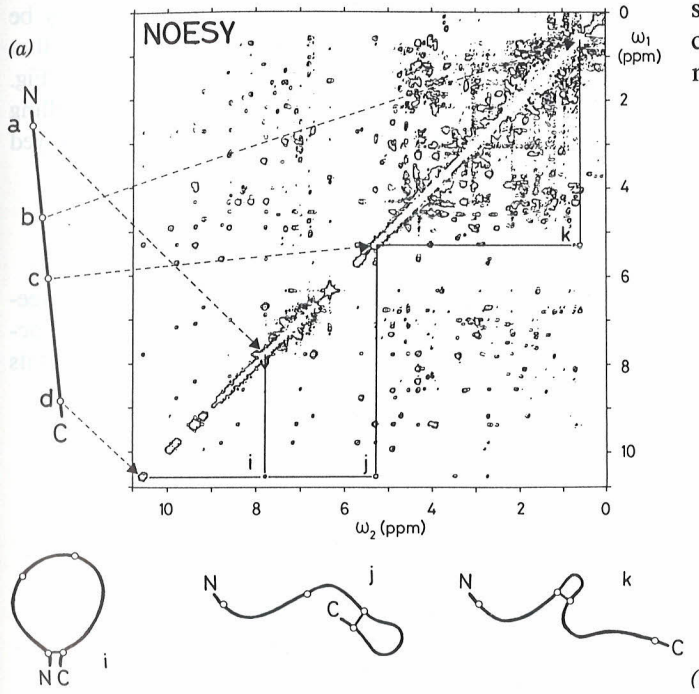


FIGURE 7-38. A section through the electron density map of diketopiperazine calculated at the indicated resolution levels. Hydrogen atoms are not apparent in this map because of their low electron density. [After Hodgkin, D.C., *Nature* **188,** 445 (1960).]

2. A protein may crystallize in one of several forms or "habits," depending on crystallization conditions, that differ in how the protein molecules are arranged in space relative to each other. In the numerous cases in which different crystal forms of the same protein have been independently analyzed, the molecules have virtually identical conformations. Similarly, in the several cases that both the X-ray crystal structure and the solution NMR structure of the same protein have been determined, the two structures are, for the most part, identical to within experimental error (see below). Evidently, crystal packing forces do not greatly perturb the structures of protein molecules.
3. The most compelling evidence that crystalline proteins have biologically relevant structures, however, is the observation that many enzymes are catalytically active in the crystalline state. The catalytic activity of an enzyme is very sensitive to the relative orientations of the groups involved in binding and catalysis (Chapter 14). Active crystalline enzymes must therefore have conformations that closely resemble their solution conformations.

Protein Structure Determination by 2D-NMR

The determination of the three-dimensional structures of small globular proteins in aqueous solution has become possible, since the mid 1980s, through the development of **two-dimensional (2D) NMR spectroscopy** (and, more recently, of 3D and 4D techniques), in large part by Kurt Wüthrich. Such NMR measurements, whose description is beyond the scope of this text, yield the interatomic distances between specific protons that are <5 Å apart in a protein of known sequence that has no more than ~200 residues. The interproton distances may be either through space, as determined by nuclear Overhauser effect spectroscopy (NOESY, Fig. 7-39*a*), or through bonds, as determined by correlated spectroscopy (COSY). These distances are schematically indicated by the three circular structures drawn below the spectrum. Note that the assignment of a distance relationship between two protons in a polypeptide requires that the NMR peaks to which they give rise and their positions in the polypeptide be known, which requires that the polypeptide's amino acid sequence has been previously determined. [After Wüthrich, K., *Science* **243,** 45 (1989).]



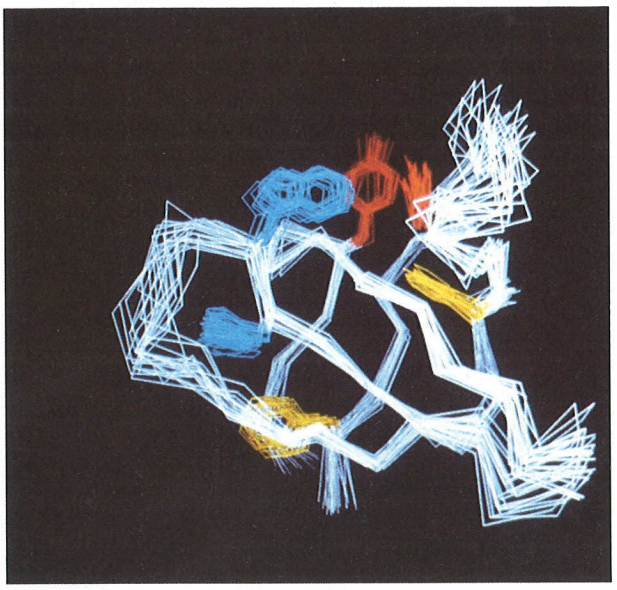


FIGURE 7-39. The 2D proton NMR structures of proteins. (*a*) A NOESY spectrum of a protein presented as a contour plot with two frequency axes, ω_1 and ω_2. The conventional 1D-NMR spectrum of the protein, which occurs along the diagonal of the plot ($\omega_1 = \omega_2$), is too crowded with peaks to be directly interpretable (even a small protein has hundreds of protons). The off-diagonal peaks, the so-called cross peaks, each arise from the interaction of two protons that are <5 Å apart in space and whose 1D-NMR peaks are located where the horizontal and vertical lines through the cross peak intersect the diagonal [a **nuclear Overhauser effect (NOE)**]. For example, the line to the left of the spectrum represents the extended polypeptide chain with its N- and C-terminal ends identified by the letters N and C and with the positions of four protons, a to d, represented by small circles. The dashed arrows indicate the diagonal NMR peaks to which these protons give rise. Cross peaks, such as *i, j,* and *k,* which are each located at the intersections of the horizontal and vertical lines through two diagonal peaks, are indicative of an NOE between the corresponding two protons, indicating that they are <5 Å apart. These distance relationships are schematically indicated by the three circular structures drawn below the spectrum. Note that the assignment of a distance relationship between two protons in a polypeptide requires that the NMR peaks to which they give rise and their positions in the polypeptide be known, which requires that the polypeptide's amino acid sequence has been previously determined. [After Wüthrich, K., *Science* **243,** 45 (1989).] (*b*) The NMR structure of a 64-residue polypeptide comprising the **Src protein SH3 domain** (Section 34-4B). The drawing represents 20 superimposed structures that are consistent with the 2D- and 3D-NMR spectra of the protein (each calculated from a different, randomly generated starting structure). The polypeptide backbone, as represented by its connected C_α atoms, is white and its Phe, Tyr, and Trp side chains are yellow, red, and blue, respectively. It can be seen that the polypeptide backbone folds into two 3-stranded antiparallel β sheets that form a sandwich. [Courtesy of Stuart Schreiber, Harvard University.]

tances, together with known geometric constraints such as covalent bond distances and angles, group planarity, chirality, and van der Waals radii, are used to compute the protein's three-dimensional structure. However, since interproton distance measurements are imprecise, they are insufficient to imply a unique structure. Rather, they are consistent with an ensemble of closely related structures. Consequently, an NMR structure of a protein (or any other macromolecule with a well-defined structure) is often presented as a representative sample of structures that are consistent with the constraints (e.g., Figure 7-39*b*). The "tightness" of a bundle of such structures is indicative both of the accuracy with which the structure is known, which in the most favorable cases is roughly comparible to that of an X-ray crystal structure with a resolution of 2 to 2.5 Å, and of the conformational fluctuations that the protein undergoes (Section 8-2).

In most of the several cases in which both the NMR and X-ray crystal structures of a particular protein have been determined, the two structures are in good agreement. There are, however, a few instances in which there are real differences between the corresponding X-ray and NMR structures. These, for the most part, involve surface residues that, in the crystal, participate in intermolecular contacts and are thereby perturbed from their solution conformations. NMR methods, besides providing mutual crosschecks with X-ray techniques, can determine the structures of proteins and other macromolecules that fail to crystallize. Moreover, since NMR can probe motions over time scales spanning 10 orders of magnitude, it can be used to study protein folding and dynamics (Chapter 8).

Protein Molecular Structures Are Most Effectively Illustrated in Simplified Form

The several hundred nonhydrogen atoms of even a small protein makes understanding the detailed structure of a protein a considerable effort. The most instructive method of studying a protein structure is the hands-on examination of its skeletal (ball-and-stick) model. Unfortunately, such models are rarely available and photographs of them are too cluttered to be of much use. A practical alternative is a computer-generated stereo diagram in which the polypeptide backbone is represented only by its C_α atoms and only a few key side chains are included (Fig. 7-40). Another possibility is an artistic rendering of a protein model that has been simplified and slightly distorted to improve its visual clarity (Fig. 7-41). A further level of abstraction may be obtained by representing the protein in a cartoon form that emphasizes its secondary structure (Fig. 7-42; also see Fig. 7-19). Computer-generated drawings of space-filling models, such as Figs. 7-12 and 7-18, may also be employed to illustrate certain features of protein structures.

B. Tertiary Structure

The **tertiary structure (3° structure)** of a protein is its three-dimensional arrangement; that is, the folding of its 2° structural elements, together with the spatial dispositions of its

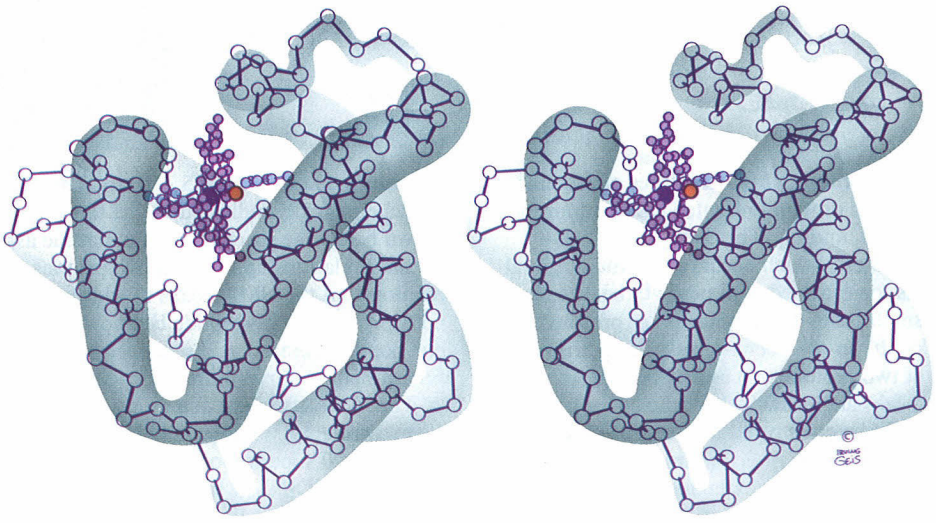


FIGURE 7-40. A computer-drawn stereo diagram of sperm whale myoglobin in which the C_α atoms are represented by balls and the peptide groups linking them are represented by solid bonds. The 153-residue polypeptide chain is folded into eight α helices (highlighted here by hand-drawn envelopes), connected by short polypeptide links. The protein's bound heme group (*purple*) in complex with an O_2 molecule (*orange sphere*) is shown together with its two closely associated His side chains (*light blue*). Hydrogen atoms have been omitted for the sake of clarity. Instructions for viewing stereo diagrams are given in the appendix to this chapter. [Figure copyrighted © by Irving Geis.]

side chains. The first protein X-ray structure, that of sperm whale myoglobin, was elucidated in the late 1950s by John Kendrew and coworkers. Its polypeptide chain follows such a tortuous, wormlike path (Figs. 7-40 through 7-42), that these investigators were moved to indicate their disappointment at its lack of regularity. In the intervening years, well over 500 protein structures have been reported. Each of them is a unique, highly complicated entity. Nevertheless, their tertiary structures have several outstanding features in common as we shall see below.

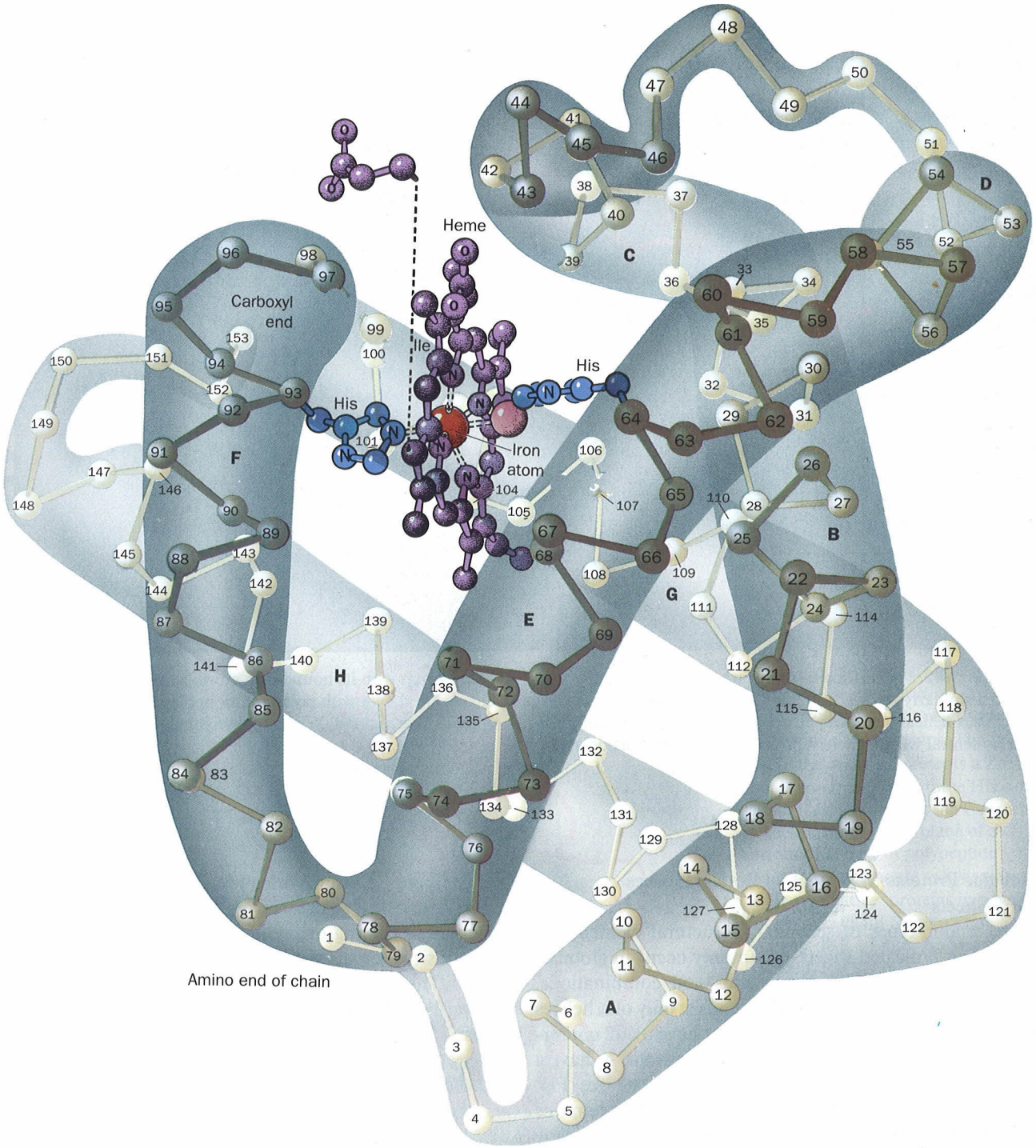


FIGURE 7-41. An artist's rendering of sperm whale myoglobin analogous to Fig. 7-40. One of the heme group's propionic acid side chains has been displaced for clarity. The amino acid residues are consecutively numbered, starting from the N-terminus, and the eight helices are likewise designated A through H. [Figure copyrighted © by Irving Geis.]

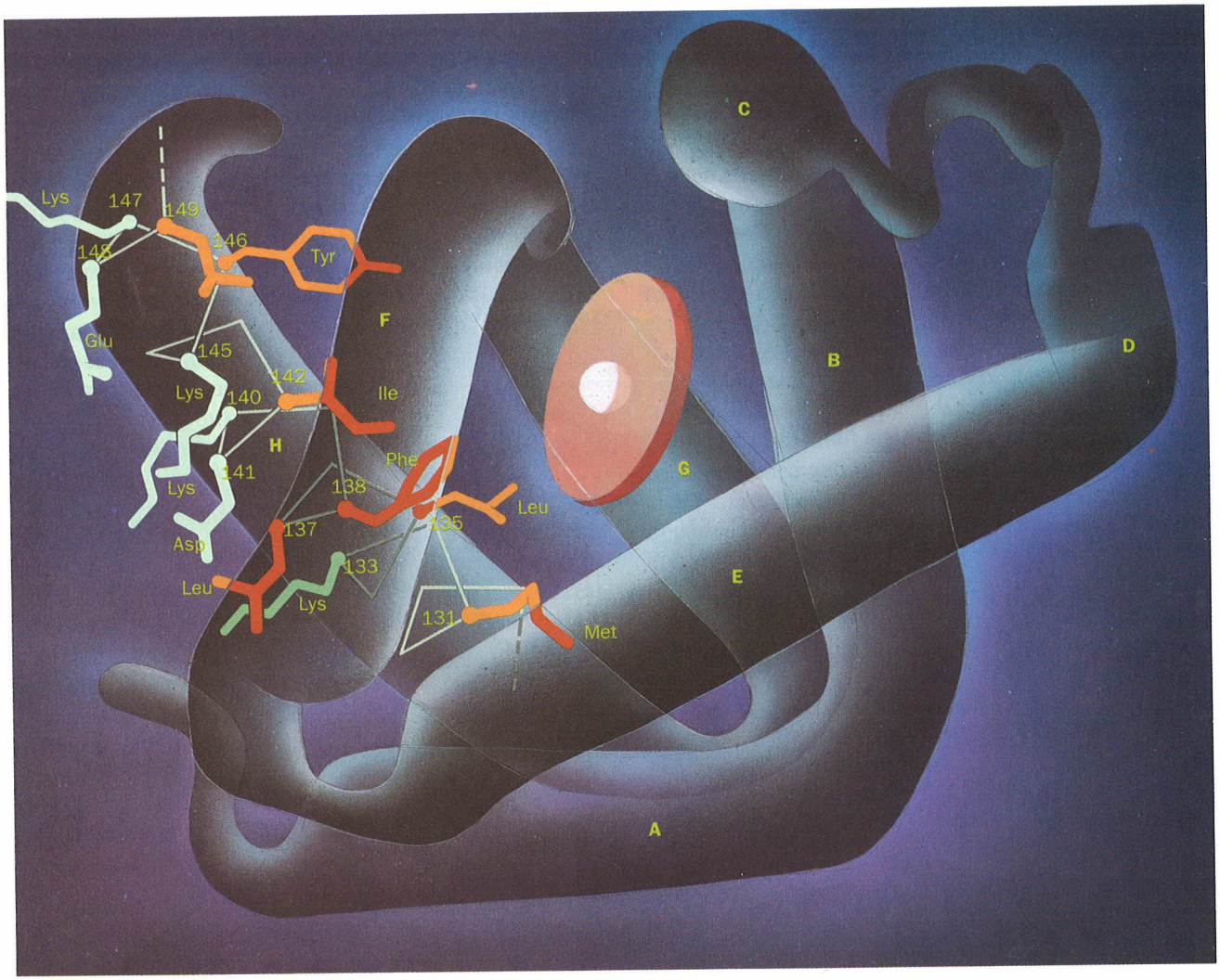


FIGURE 7-42. A cartoon of sperm whale myoglobin, oriented similarly to Figs. 7-40 and 7-41, which emphasizes the protein's α helical secondary structure (*cylinders*). The pink disk with its central white sphere represents the protein's associated heme group with its bound iron atom. Many of the H helix side chains are shown with polar and nonpolar groups, respectively, colored blue and red. [Figure copyrighted © by Irving Geis.]

Globular Proteins May Contain Both α Helices and β Sheets

The major types of secondary structural elements, α helices and β pleated sheets, commonly occur in globular proteins but in varying proportions and combinations. Some proteins, such as myoglobin, consist only of α helices spanned by short connecting links that have a coil conformation (Fig. 7-42). Others, such as concanavalin A, have a large proportion of β sheets but are devoid of α helices (Fig. 7-43). Most proteins, however, have significant amounts of both types of secondary structure (on average, ~31% α helix and ~28% β sheet). Human **carbonic anhydrase** (Fig. 7-44) as well as carboxypeptidase and triose phosphate isomerase (Fig. 7-19) are examples of such proteins.

Side Chain Location Varies with Polarity

The primary structures of globular proteins generally lack the repeating or pseudorepeating sequences that are responsible for the regular conformations of fibrous proteins. The amino acid side chains in globular proteins are, nevertheless, spatially distributed according to their polarities:

1. *The nonpolar residues Val, Leu, Ile, Met, and Phe largely occur in the interior of a protein, out of contact with the aqueous solvent.* The hydrophobic interactions that promote this distribution, which are largely responsible for the three-dimensional structures of native proteins, are further discussed in Section 7-4C.

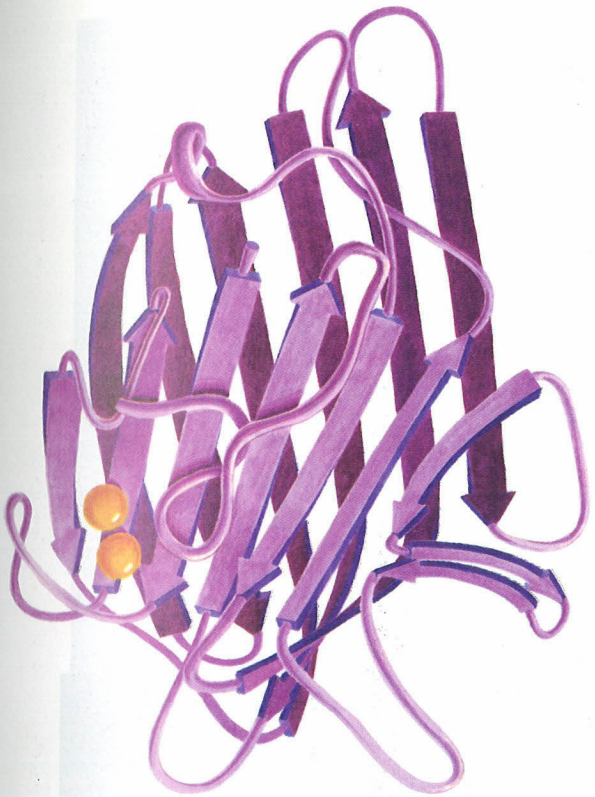
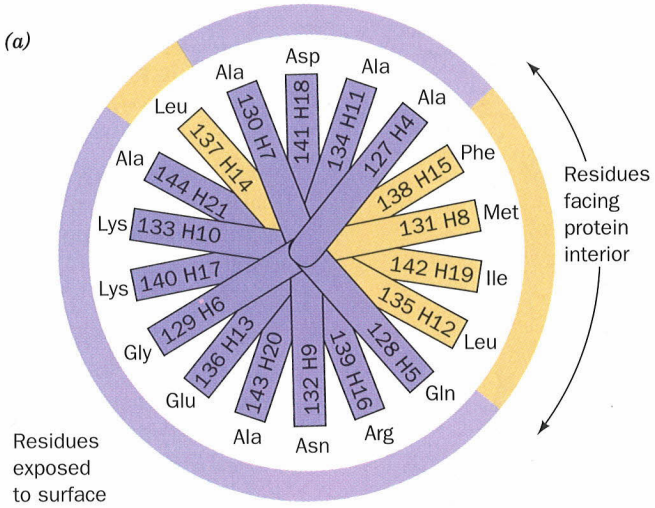
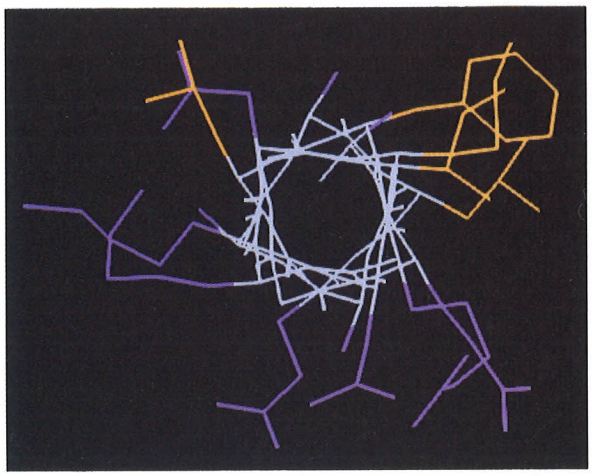


FIGURE 7-43. The jack bean protein concanavalin A largely consists of extensive regions of antiparallel β pleated sheet, here represented by arrows pointing towards the polypeptide chain's C-terminus. The balls represent protein-bound metal ions. The back sheet is shown in a space-filling representation in Fig. 7-18. [After a drawing by Jane Richardson, Duke University.]

2. *The charged polar residues Arg, His, Lys, Asp, and Glu are largely located on the surface of a protein in contact with the aqueous solvent.* This is because the immersion of an ion in the virtually anhydrous interior of a protein results in the uncompensated loss of much of its hydration energy. In the instances that these groups occur in the interior of a protein, they often have a specific chemical function such as promoting catalysis or participating in metal ion binding (e.g., the metal ion–liganding His residues in Figs. 7-41 and 7-44).
3. The uncharged polar groups Ser, Thr, Asn, Gln, Tyr, and Trp, are usually on the protein surface but frequently occur in the interior of the molecule. In the latter case, these residues are almost always hydrogen bonded to other groups in the protein. In fact, *nearly all buried hydrogen bond donors form hydrogen bonds with buried acceptor groups;* in a sense, the formation of a hydrogen bond "neutralizes" the polarity of a hydrogen bonding group.

FIGURE 7-44. Human carbonic anhydrase in which α helices are represented as cylinders and each strand of β sheet is drawn as an arrow pointing towards the polypeptide's C-terminus. The gray ball in the middle represents a Zn^{2+} ion that is coordinated by three His side chains (*blue*). Note that the C-terminus is tucked through the plane of a surrounding loop of polypeptide chain so that carbonic anhydrase is one of the rare native proteins in which a polypeptide chain forms a knot. [After Kannan, K.K., Liljas, A., Waara, I., Bergsten, P.-C., Lovgren, S., Strandberg, B., Bengtsson, J., Carlbom, U., Friedborg, K., Jarup, L., and Petef, M., *Cold Spring Harbor Symp. Quant. Biol.* **36,** 221 (1971).]





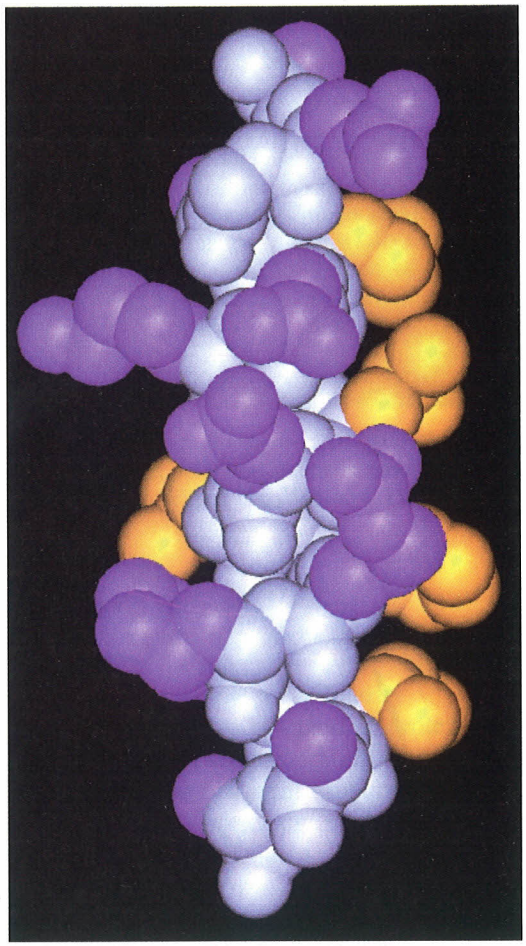


FIGURE 7-45. The H helix of sperm whale myoglobin. (*a*) A **helical wheel** representation in which side chain positions about the α helix are projected down the helix axis onto a plane. Here each residue is identified both according to its sequence in the polypeptide chain, and according to its position in the H helix. The residues lining the side of the helix facing the protein's interior regions are all nonpolar. The other residues, except Leu 137, which contacts the protein segment linking helices E and F (Figs. 7-40 and 7-41), are exposed to the solvent and are all more or less polar. (*b*) A skeletal model, viewed as in Part *a*, in which the main chain is white, nonpolar side chains are yellow, and polar side chains are purple. (*c*) A space-filling model, viewed from the bottom of the page in Parts *a* and *b* and colored as in Part *b*. Compare these diagrams with the drawing of the H helix in Fig. 7-42.

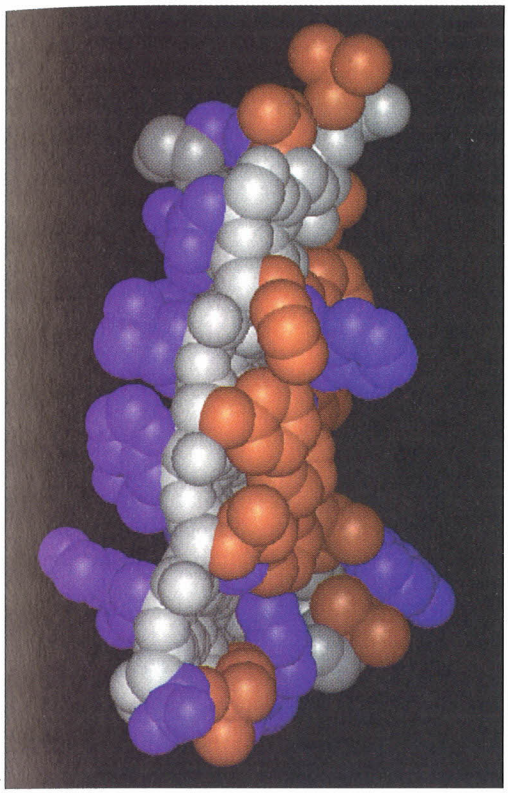


FIGURE 7-46. A space-filling model of an antiparallel β sheet from concanavilin A in side view with the interior of the protein (the surface of a second antiparallel β sheet; see Fig. 7-43) to the right and the exterior to the left. The main chain is white, nonpolar side chains are brown, and polar side chains are purple.

This side chain distribution is apparent in Figs. 7-42 and 7-45, which show the surface and interior exposures of the amino acid side chains of myoglobin's H helix. This arrangement is likewise seen on the covers of this textbook, which show the distributions of polar *(front cover)* and nonpolar *(back cover)* residues of cytochrome *c,* as well as in Fig. 7-46, which shows one of the antiparallel β pleated sheets of concanavilin A.

Globular Protein Cores Are Efficiently Arranged With Their Side Chains in Relaxed Conformations

Globular proteins are quite compact; there is very little space inside them so that water is largely excluded from their interiors. The micellelike arrangement of their side chains (polar groups outside, nonpolar groups inside) has led to their description as "oil drops with polar coats." This generalization, although picturesque, lacks precision. The **packing density** (ratio of the volume enclosed by the van der Waals envelopes of the atoms in a region to the total volume of the region) of the internal regions of globular proteins averages ~0.75, which is in the same range as that of molecular crystals of small organic molecules. In comparison, equal-sized close-packed spheres have a packing density of 0.74, whereas organic liquids (oil drops) have packing densities that are mostly between 0.60 and 0.70. *The interior of a protein is therefore more like a molecular crystal than an oil drop; that is, it is efficiently packed.* The ability of most hydrogen bonding donors to find acceptors under such constrained conditions is explained by the observation that most hydrogen bonding partners reside on residues that are close in sequence (which is, in turn, explained by the facts that backbone N—H groups comprise the majority of the hydrogen bonding donors in proteins and that most protein residues are members of secondary structural elements).

The bonds of protein side chains, including those occupying protein cores, almost invariably have low-energy staggered torsion angles (Fig. 7-5*b*). Evidently, interior side chains adopt relaxed conformations despite their profusion of intramolecular interactions (Section 7-4).

Large Polypeptides Form Domains

Polypeptide chains that consist of more than ~200 residues usually fold into two or more globular clusters known as **domains,** which give these proteins a bi- or multilobal appearance. Most domains consist of 100 to 200 amino acid residues and have an average diameter of ~25 Å. Each subunit of **glyceraldehyde-3-phosphate dehydrogenase,** for example, has two distinct domains (Fig. 7-47). A polypeptide chain wanders back and forth within a domain but neighboring domains are usually connected by one, or less commonly two, polypeptide segments. *Domains are therefore structurally independent units that each have the characteristics of a small globular protein.* Indeed, limited proteolysis of a multidomain protein often liberates its domains without greatly altering their structures. Nevertheless, the domain structure of a protein is not always obvious since its domains may make such extensive contacts with each other that the protein appears to be a single globular entity.

An inspection of the various protein structures diagrammed in this chapter reveals that domains consist of two or more layers of secondary structural elements. The reason for this is clear: At least two such layers are required to seal off a domain's hydrophobic core from the aqueous environment.

Domains often have a specific function such as the binding of a small molecule. In Fig. 7-47, for example, **nicotinamide adenine dinucleotide (NAD^+)** binds to the first domain of glyceraldehyde-3-phosphate dehydrogenase. Small molecule–binding sites in multidomain proteins often occur in the clefts between domains; that is, the small molecules are bound by groups from two domains. This arrangement arises, in part, from the need for a flexible interaction between the protein and the small molecule that the relatively pliant covalent connection between the domains can provide.

Supersecondary Structures Have Structural and Functional Roles

Certain groupings of secondary structural elements, named **supersecondary structures** or **motifs,** occur in many unrelated globular proteins:

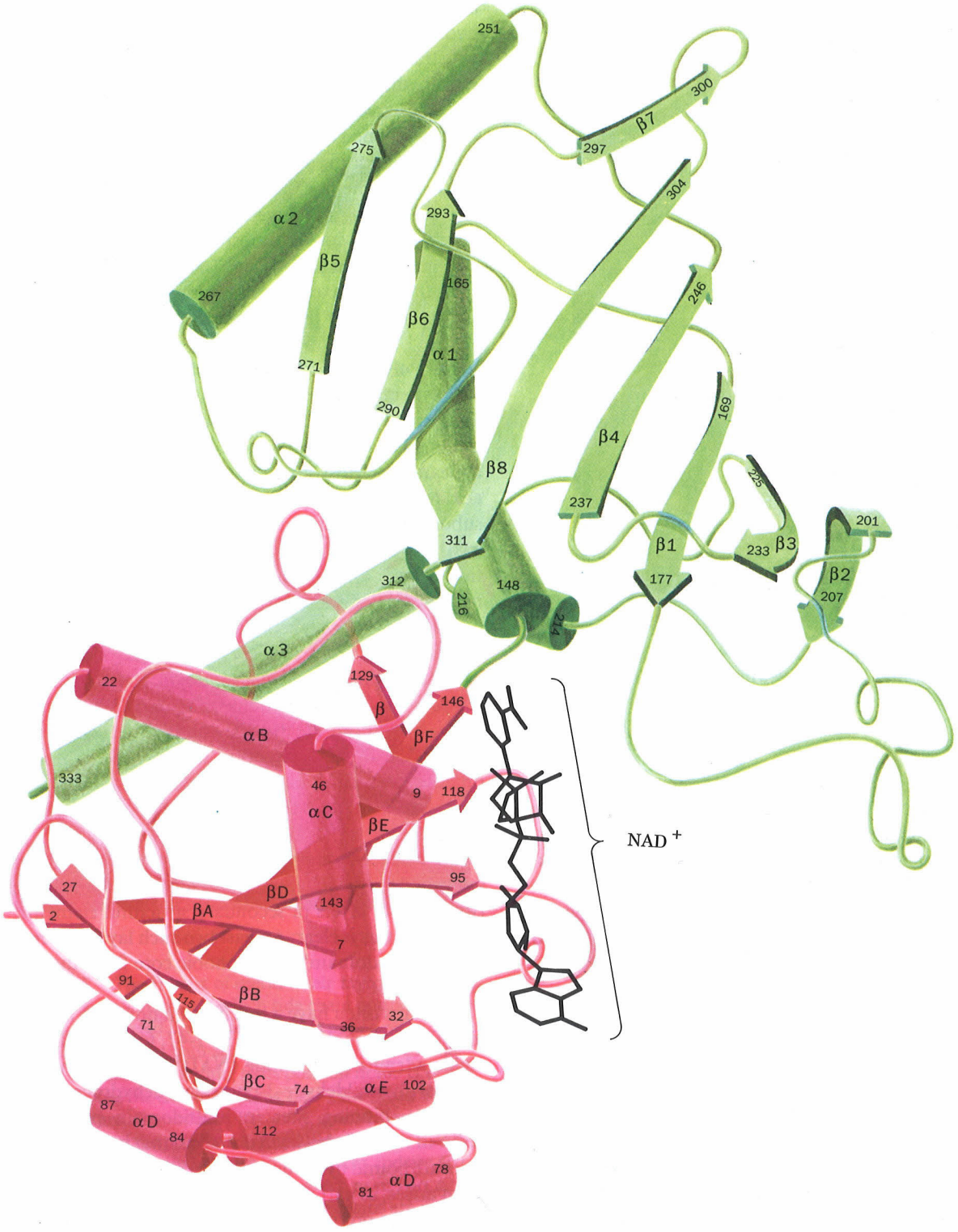
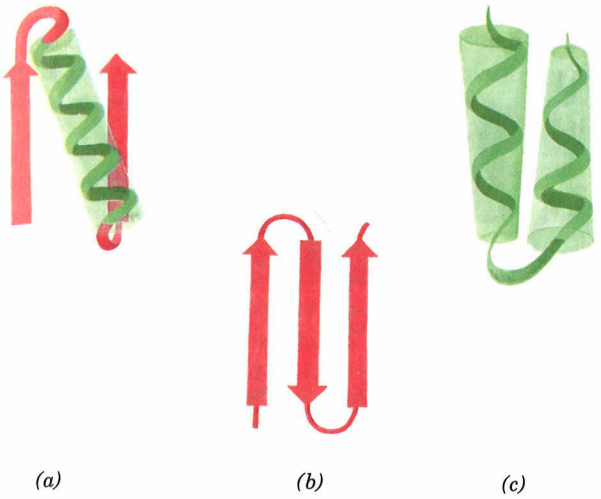


FIGURE 7-47. One subunit of the enzyme glyceraldehyde-3-phosphate dehydrogenase from *Bacillus stearothermophilus.* The polypeptide folds into two distinct domains. The first domain (*red,* residues 1–146) binds NAD^+ (*black*) near the C-terminal ends of its parallel β strands, and the second domain (*green*) binds glyceraldehyde-3-phosphate (not shown). [After Biesecker, G., Harris, J.I., Thierry, J.C., Walker, J.E., and Wonacott, A., *Nature* **266,** 331 (1977).]

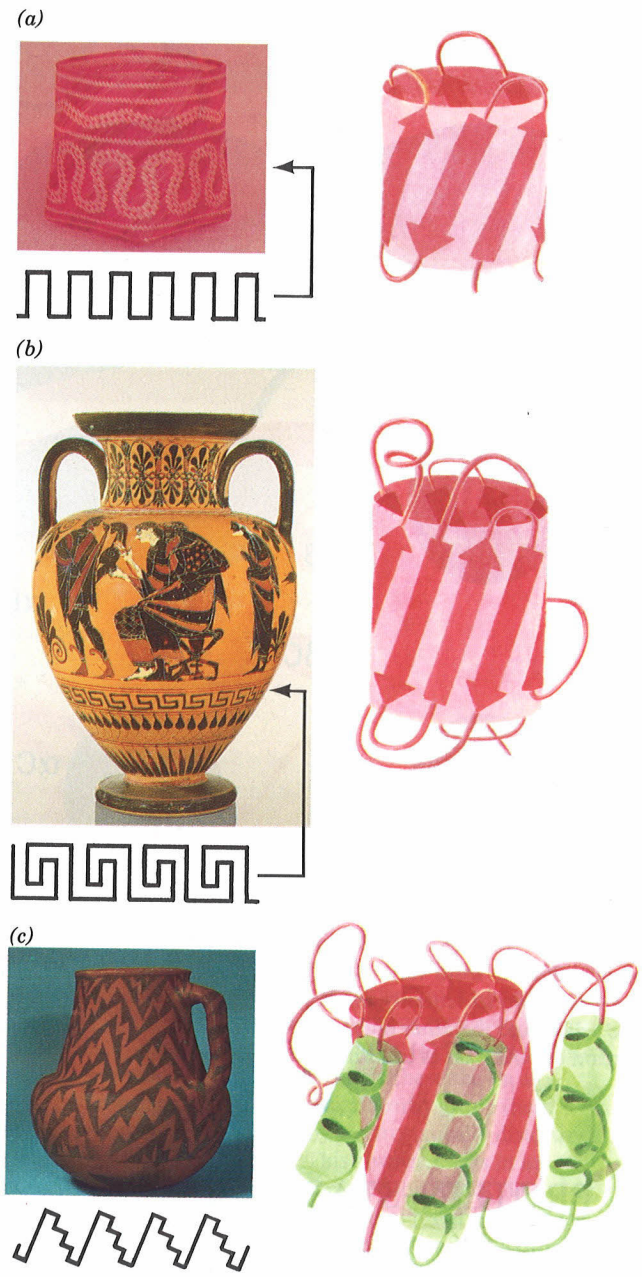


(a) *(b)* *(c)*

FIGURE 7-48. Schematic diagrams of (*a*) a $\beta\alpha\beta$ motif, (*b*) a β hairpin motif, and (*c*) an $\alpha\alpha$ motif.

1. The most common form of supersecondary structure is the **$\beta\alpha\beta$ motif,** in which the usually right-handed crossover connection between two consecutive parallel strands of a β sheet consists of an α helix (Fig. 7-48*a*).
2. Another common supersecondary structure, the **β hairpin motif,** consists of an antiparallel β sheet formed by sequential segments of polypeptide chain that are connected by relatively tight reverse turns (Fig. 7-48*b*).
3. In an **$\alpha\alpha$ motif**, two successive antiparallel α helices pack against each other with their axes inclined so as to permit energetically favorable intermeshing of their contacting side chains (Fig. 7-48*c*). Such associations stabilize the coiled coil conformation of α keratin (Section 7-2A).
4. Extended β sheets often roll up to form **β barrels** (e.g., Fig. 7-19*b*). Three different β barrel topologies (the ways in which the strands and their interconnections are arranged) have been named in analogy with geometric motifs found on Native American and Greek weaving and pottery (Fig. 7-49).

Supersecondary structures may have functional as well as structural significance. A $\beta\alpha\beta\alpha\beta$ unit, for example, in which the β strands form a parallel sheet with right-handed α helical crossover connections (two overlapping $\beta\alpha\beta$ units), was shown by Michael Rossmann to form a nucleotide-binding site in many enzymes. In most proteins that bind dinucleotides, two such $\beta\alpha\beta\alpha\beta$ units combine to form a motif alternatively known as a **dinucleotide-binding fold** or a **Rossmann fold** (Fig. 7-50). In some cases, the second α helix in a $\beta\alpha\beta\alpha\beta$ unit is replaced by a length of nonhelical polypeptide. This occurs, for example, between the βE and βF strands of glyceraldehyde-3-phosphate dehydrogenase (Fig. 7-47).



(a) *(b)* *(c)*

FIGURE 7-49. Comparisons of the backbone folding patterns of protein β barrels (*right*) with geometric motifs commonly used to decorate Native American and Greek weaving and pottery (*left*). (*a*) Native American polychrome cane basket and the polypeptide backbone of **rubredoxin** from *Clostridium pasteurianum* showing its linked β meanders. [Museum of the American Indian, Heye Foundation.] (*b*) Red figured Greek amphora with its Greek key border area showing Cassandra and Ajax (about 450 B.C.) and the polypeptide backbone of human **prealbumin** with its "Greek key" pattern. [The Metropolitan Museum of Art, Fletcher Fund, 1956.] (*c*) Early Anasazi redware pitcher from New Mexico and the polypeptide backbone of chicken muscle triose phosphate isomerase showing its "lightning" pattern of overlapping $\beta\alpha\beta$ units. This so-called α/β barrel is also diagrammed in Fig. 7-19*b*. [Museum of the American Indian, Heye Foundation.] [After Richardson, J.S., *Nature* **268,** 498 (1977).]

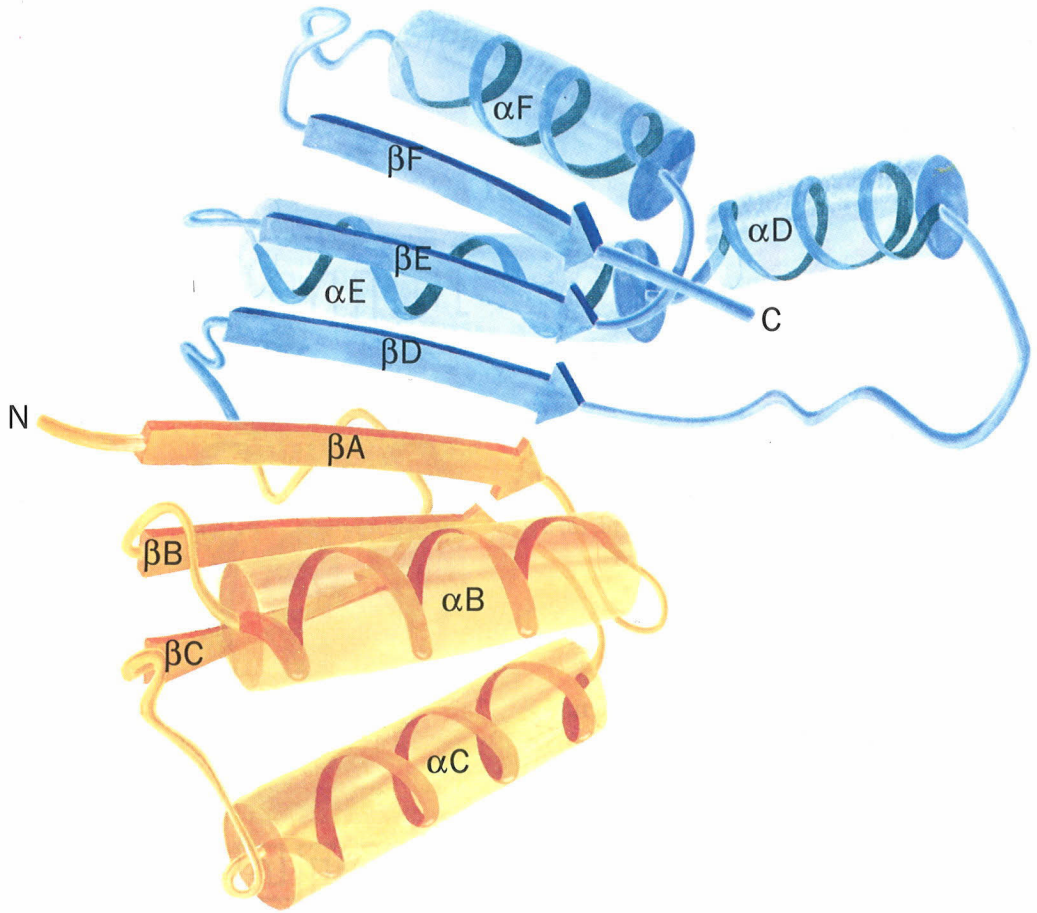


FIGURE 7-50. An idealized representation of the coenzyme-binding domain from various dehydrogenases. This domain consists of two structurally similar *βαβαβ* units, drawn here with one yellow and the other blue, each of which binds a nucleotide portion of NAD^+ so as to form a dinucleotide-binding or Rossmann fold. Compare this figure with the NAD^+-binding domain of glyceraldehyde-3-phosphate dehydrogenase (Fig. 7-47). [After Rossmann, M.G., Liljas, A., Bränden, C.-I., and Banaszak, L.J., *in* Boyer, P.D. (Ed.), *The Enzymes,* Vol. 11 (3rd ed.), *p.* 68, Academic Press (1975).]

4. PROTEIN STABILITY

Incredible as it may seem, thermodynamic measurements indicate that *native proteins are only marginally stable entities under physiological conditions.* The free energy required to denature them is ~0.4 $kJ \cdot mol^{-1}$ of amino acid residues so that 100-residue proteins are typically stable by only around 40 $kJ \cdot mol^{-1}$. In contrast, the energy required to break a typical hydrogen bond is ~20 $kJ \cdot mol^{-1}$. The various noncovalent influences to which proteins are subject—electrostatic interactions (both attractive and repulsive), hydrogen bonding (both intramolecular and to water), and hydrophobic forces—each have energetic magnitudes that may total thousands of kilojoules per mole over an entire protein molecule. Consequently, *a protein structure is the result of a delicate balance among powerful countervailing forces.* In this section we discuss the nature of these forces and end by considering protein denaturation: that is, how these forces can be disrupted.

A. Electrostatic Forces

Molecules are collections of electrically charged particles and hence, to a reasonable degree of approximation, their interactions are determined by the laws of classical electrostatics (more exact calculations require the application of quantum mechanics). The energy of association, U, of two electric charges, q_1 and q_2, that are separated by the distance r, is found by integrating the expression for Coulomb's law, Eq. [2.1], to determine the work necessary to separate these charges by an infinite distance:

$$U = \frac{kq_1q_2}{Dr} \qquad [7.1]$$

Here $k = 9.0 \times 10^9\ J \cdot m \cdot C^{-2}$ and D is the dielectric constant of the medium in which the charges are immersed (recall that $D = 1$ for a vacuum and, for the most part, increases with the polarity of the medium; Table 2-1). The dielectric constant of a molecule-sized region is difficult to

estimate. For the interior of a protein, it is usually taken to be in the range 3 to 5 in analogy with the measured dielectric constants of substances that have similar polarities such as benzene and diethyl ether.

Ionic Interactions Are Strong but Do Not Greatly Stabilize Proteins

The association of two ionic protein groups of opposite charge is known as an **ion pair** or **salt bridge.** According to Eq. [7.1], the energy of a typical ion pair, say the carboxyl group of Glu and the ammonium group of Lys, whose charge centers are separated by 4.0 Å in a medium of dielectric constant 4, is $-86\ kJ \cdot mol^{-1}$ (one electronic charge = 1.60×10^{-19} C). Free ions in aqueous solution are highly solvated, however, so that the free energy of solvation of two separated ions is about equal to the free energy of formation of their unsolvated ion pairs. *Ion pairs therefore contribute little stability towards a protein's native structure.* This accounts for the observations that although ~75% of charged residues occur in ion pairs, very few ion pairs are buried (unsolvated) and that ion pairs that are exposed to the aqueous solvent tend to be but poorly conserved among homologous proteins.

Dipole–Dipole Interactions Are Weak but Significantly Stabilize Protein Structures

The noncovalent associations between electrically neutral molecules, collectively known as **van der Waals forces,** arise from electrostatic interactions among permanent and/or induced dipoles. These forces are responsible for numerous interactions of varying strengths between nonbonded neighboring atoms. (The hydrogen bond, a special class of dipolar interaction, is considered separately in Section 7-4B.)

Interactions among permanent dipoles are important structural determinants in proteins because many of their groups, such as the carbonyl and amide groups of the peptide backbone, have permanent dipole moments. These interactions are generally much weaker than the charge–charge interactions of ion pairs. Two carbonyl groups, for example, each with dipoles of 4.2×10^{-30} C·m (1.3 debye units) that are oriented in an optimal head-to-tail arrangement (Fig. 7-51*a*) and separated by 5 Å in a medium of dielectric constant 4, have a calculated attractive energy of only $-9.3\ kJ \cdot mol^{-1}$. Furthermore, these energies vary with r^{-3} so they rapidly attenuate with distance. In α helices, however, the dipolar amide and carbonyl groups of the polypeptide backbone all point in the same direction (Fig. 7-11) so that their interactions are associative and tend to be additive (these groups, of course, also form hydrogen bonds but here we are concerned with their residual electric fields). The carbonyl groups all have their oxygen atoms pointing towards the C terminal end of the α helix, giving it a significant dipole moment that is positive towards the N terminus and negative towards the C terminus. Consequently, *in the low dielectric constant core of a protein, dipole–dipole interactions significantly influence protein folding.*

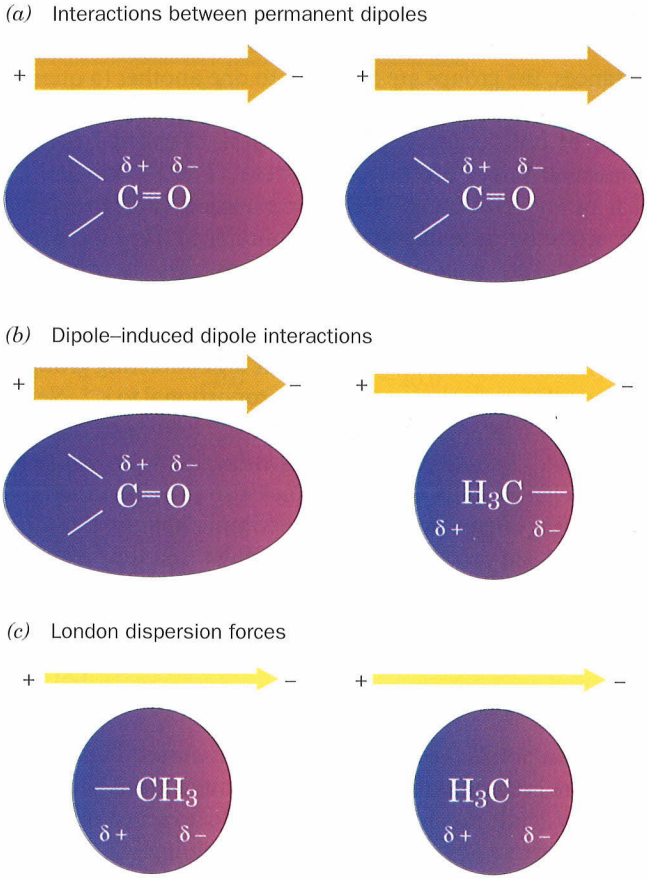


FIGURE 7-51. Dipole–dipole interactions. The strength of each dipole is represented by the thickness of the accompanying arrow. (*a*) Interactions between permanent dipoles. These interactions, here represented by carbonyl groups lined up head to tail, may be attractive, as shown here, or repulsive, depending on the relative orientations of the dipoles. (*b*) Dipole–induced dipole interactions. A permanent dipole (here shown as a carbonyl group) induces a dipole in a nearby group (here represented by a methyl group) by electrostatically distorting its electron distribution (*shading*). This always results in an attractive interaction. (*c*) London dispersion forces. The instantaneous charge imbalance (*shading*) resulting from the motions of the electrons in a molecule (*left*) induce a dipole in a nearby group (*right*); that is, the motions of the electrons in neighboring groups are correlated. This always results in an attractive interaction.

A permanent dipole also induces a dipole moment on a neighboring group so as to form an attractive interaction (Fig. 7-51*b*). Such dipole–induced dipole interactions are generally much weaker than are dipole–dipole interactions.

Although nonpolar molecules are nearly electrically neutral, at any instant they have a small dipole moment resulting from the rapid fluctuating motion of their electrons. This transient dipole moment polarizes the electrons in a

neighboring group, thereby giving rise to a dipole moment (Fig. 7-51*c*) such that, near their van der Waals contact distances, the groups are attracted to one another (a quantum mechanical effect that really cannot be explained in terms of only classical physics). These so-called **London dispersion forces** are extremely weak. The 8.2-kJ·mol^{-1} heat of vaporization of CH_4, for example, indicates that the attractive interaction of a nonbonded H···H contact between neighboring CH_4 molecules is roughly -0.3 kJ·mol^{-1} (in the liquid, a CH_4 molecule touches its 12 nearest neighbors with ~2 H···H contacts each).

London forces are only significant for contacting groups because their association energy is proportional to r^{-6}. Nevertheless, *the great numbers of interatomic contacts in proteins makes London forces a major influence in determining their conformations.* London forces also provide much of the binding energy in the sterically complementary interactions between proteins and the molecules that they specifically bind.

B. Hydrogen Bonding Forces

Hydrogen bonds (D—H···A), as we discussed in Section 2-1A, are predominantly electrostatic interactions between a weakly acidic donor group (D—H) and an acceptor atom (A) that bears a lone pair of electrons. In biological systems, D and A can both be the highly electronegative N and O atoms and occasionally S atoms. Hydrogen bonds, which have association energies in the range -12 to -30 kJ·mol^{-1}, are much more directional than are van der Waals forces although less so than are covalent bonds. The D···A distance is normally in the range 2.7 to 3.1 Å. Hydrogen bonds tend to be linear with the D—H bond pointing along the acceptor's lone pair orbital. Large deviations from this ideal geometry are not unusual, however. For example, in the hydrogen bonds of both α helices (Fig. 7-11) and antiparallel β pleated sheets (Fig. 7-16*a*), the N—H bonds point approximately along the C=O bonds rather than along an O lone pair orbital, and in parallel β pleated sheets (Fig. 7-16*b*), the hydrogen bonds depart significantly from linearity. Indeed, many of the hydrogen bonds in proteins are members of networks in which each donor is hydrogen bonded to multiple acceptors and each acceptor is hydrogen bonded to multiple donors.

The internal hydrogen bonding groups of a protein are arranged such that nearly all possible hydrogen bonds are formed (Section 7-3B). Clearly, hydrogen bonding has a major influence on the structures of proteins. An unfolded protein, however, makes all its hydrogen bonds with the water molecules of the aqueous solvent (water, it will be recalled, is a strong hydrogen bonding donor and acceptor). The free energy of stabilization that internal hydrogen bonds confer upon a native protein is therefore equal to the difference in the free energy of hydrogen bonding between the native protein and the unfolded protein. Since the various hydrogen bonds in question, to a first approximation, all have the same free energy, *internal hydrogen bonding cannot significantly stabilize, and, indeed, may even slightly destabilize, the structure of a native protein relative to its unfolded state.*

Despite the foregoing, *the internal hydrogen bonds of a protein provide a structural basis for its native folding pattern:* If a protein folded in a way that prevented some of its internal hydrogen bonds from forming, their free energy would be lost and such conformations would be less stable than those that are fully hydrogen bonded. Indeed, the formation of α helices and β sheets efficiently satisfies the polypeptide backbone's hydrogen bonding requirements. This argument also applies to the van der Waals forces discussed in the previous section.

C. Hydrophobic Forces

*The **hydrophobic effect** is the name given to those influences that cause nonpolar substances to minimize their contacts with water, and amphipathic molecules, such as soaps and detergents, to form micelles in aqueous solutions (Section 2-1B).* Since native proteins form a sort of intramolecular micelle in which their nonpolar side chains are largely out of contact with the aqueous solvent, *hydrophobic interactions must be an important determinant of protein structures.*

The hydrophobic effect derives from the special properties of water as a solvent, only one of which is its high dielectric constant. In fact, other polar solvents, such as dimethylsulfoxide (DMSO) and *N*,*N*-dimethylformamide (DMF), tend to denature proteins. The thermodynamic data of Table 7-4 provide considerable insight as to the origin of the hydrophobic effect because the transfer of a hydrocarbon from water to a nonpolar solvent resembles the transfer of a nonpolar side chain from the exterior of a protein in aqueous solution to its interior. The isothermal Gibbs free energy changes ($\Delta G = \Delta H - T\Delta S$) for the transfer of a hydrocarbon from an aqueous solution to a nonpolar solvent is negative in all cases, which indicates, as we know to be the case, that such transfers are spontaneous processes (oil and water do not mix). What is perhaps unexpected is that these transfer processes are endothermic (positive ΔH) for aliphatic compounds and athermic ($\Delta H = 0$) for aromatic compounds; that is, *it is enthalpically more or equally favorable for nonpolar molecules to dissolve in water than in nonpolar media.* In contrast, the entropy component of the unitary free energy change, $-T\Delta S_u$ (see footnote *a* to Table 7-4), is large and negative in all cases. Clearly, *the transfer of a hydrocarbon from an aqueous medium to a nonpolar medium is entropically driven. The same is true of the transfer of a nonpolar protein group from an aqueous environment to the protein's nonpolar interior.*

TABLE 7-4. THERMODYNAMIC CHANGES FOR TRANSFERRING HYDROCARBONS FROM WATER TO NONPOLAR SOLVENTS AT 25°C[a]

Process	ΔH ($kJ \cdot mol^{-1}$)	$-T\Delta S_u$ ($kJ \cdot mol^{-1}$)	ΔG_u ($kJ \cdot mol^{-1}$)
CH_4 in $H_2O \rightleftharpoons CH_4$ in C_6H_6	11.7	−22.6	−10.9
CH_4 in $H_2O \rightleftharpoons CH_4$ in CCl_4	10.5	−22.6	−12.1
C_2H_6 in $H_2O \rightleftharpoons C_2H_6$ in benzene	9.2	−25.1	−15.9
C_2H_4 in $H_2O \rightleftharpoons C_2H_4$ in benzene	6.7	−18.8	−12.1
C_2H_2 in $H_2O \rightleftharpoons C_2H_2$ in benzene	0.8	−8.8	−8.0
Benzene in $H_2O \rightleftharpoons$ liquid benzene[b]	0.0	−17.2	−17.2
Toluene in $H_2O \rightleftharpoons$ liquid toluene[b]	0.0	−20.0	−20.0

[a] ΔG_u, the **unitary Gibbs free energy change,** is the Gibbs free energy change, ΔG, corrected for its concentration dependence so that it reflects only the inherent properties of the substance in question and its interaction with solvent. This relationship, according to Equation [3.13], is

$$\Delta G_u = \Delta G - nRT \ln \frac{[A_f]}{[A_i]}$$

where $[A_i]$ and $[A_f]$ are the initial and final concentrations of the substance under consideration, respectively, and n is the number of moles of that substance. Since the second term in this equation is a purely entropic term (concentrating a substance increases its order), ΔS_u, the **unitary entropy change,** is expressed

$$\Delta S_u = \Delta S + nR \ln \frac{[A_f]}{[A_i]}$$

[b] Data measured at 18°C.

Source: Kauzmann, W., *Adv. Protein Chem.* **14,** 39 (1959).

What is the physical mechanism whereby nonpolar entities are excluded from aqueous solution? Recall that entropy is a measure of the order of a system; it decreases with increasing order (Section 3-2). Thus the decrease in entropy when a nonpolar molecule or side chain is solvated by water (the reverse of the foregoing process) must be due to an ordering process. This is an experimental observation, not a theoretical conclusion. The magnitudes of the entropy changes are too large to be attributed only to changes in the conformations of the hydrocarbons; rather, as Henry Frank and Marjorie Evans pointed out in 1945, *these entropy changes mainly arise from some sort of ordering of the water structure.*

Liquid water has a highly ordered and extensively hydrogen bonded structure (Section 2-1A). The insinuation of a nonpolar group into this structure disrupts it: A nonpolar group can neither accept nor donate hydrogen bonds, so the water molecules at the surface of the cavity occupied by the nonpolar group cannot hydrogen bond to other molecules in their usual fashion. In order to recover the lost hydrogen bonding energy, these surface waters must orient themselves so as to form a hydrogen bonded network enclosing the cavity (Fig. 7-52). This orientation constitutes an ordering of the water structure since the number of ways that

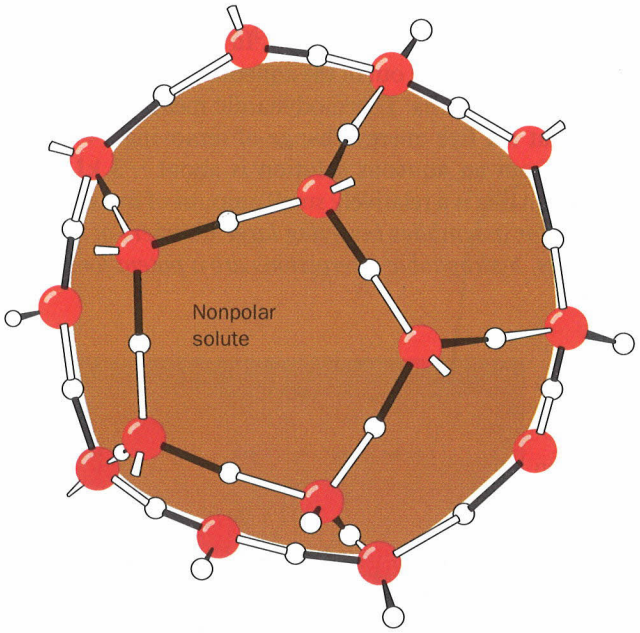


FIGURE 7-52. The orientational preference of water molecules next to a nonpolar solute. In order to maximize their hydrogen bonding energy, these water molecules tend to straddle the inert solute such that two or three of their tetrahedral directions are tangential to its surface. This permits them to form hydrogen bonds with neighboring water molecules lining the nonpolar surface. This ordering of water molecules extends several layers of water molecules beyond the first hydration shell of the nonpolar solute.

water molecules can form hydrogen bonds about the surface of a nonpolar group is less than the number of ways that they can hydrogen bond in bulk water.

Unfortunately, the complexity of liquid water's basic structure (Section 2-1A) has not yet allowed a detailed structural description of this ordering process. One model that has been proposed is that water forms quasi-crystalline hydrogen bonded cages about the nonpolar groups similar to those of **clathrates** (Fig. 7-53). The magnitudes of the entropy changes that result when nonpolar substances are dissolved in water, however, indicate that the resulting water structures can only be slightly more ordered than bulk water. They also must be quite different from that of ordinary ice, because, for instance, the solvation of nonpolar groups by water causes a large decrease in water volume (e.g., the transfer of CH_4 from hexane to water shrinks the water solution by 22.7 mL · mol^{-1} of CH_4), whereas the freezing of water results in a 1.6-mL · mol^{-1} expansion.

The unfavorable free energy of hydration of a nonpolar substance caused by its ordering of the surrounding water molecules has the net result that *the nonpolar substance is excluded from the aqueous phase.* This is because the surface area of a cavity containing an aggregate of nonpolar molecules is less than the sum of the surface areas of the cavities that each of these molecules would individually occupy. The aggregation of the nonpolar groups thereby minimizes the surface area of the cavity and therefore the entropy loss of the entire system. In a sense, the nonpolar groups are squeezed out of the aqueous phase by the hydrophobic interactions. Thermodynamic measurements indicate that the free energy change of removing a $—CH_2—$ group from an aqueous solution is about -3 kJ · mol^{-1}. Although this is a relatively small amount of free energy, *in molecular assemblies involving large numbers of nonpolar contacts, hydrophobic interactions are a potent force.*

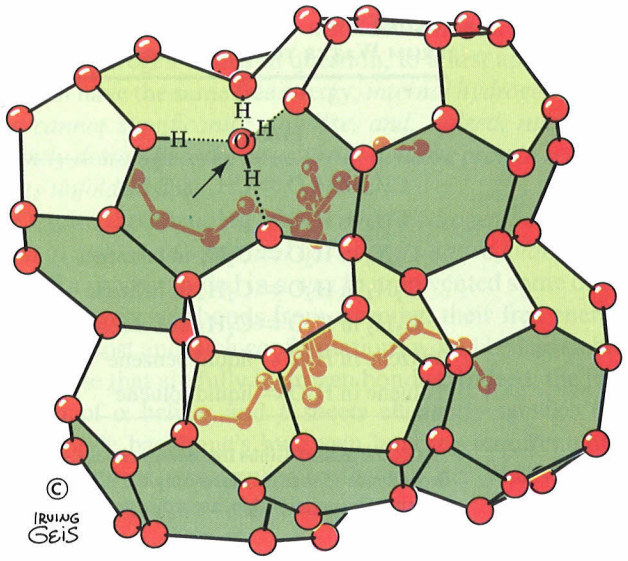


FIGURE 7-53. The structure of the clathrate $(n\text{-}C_4H_9)_3S^+F^- \cdot 23H_2O$. Clathrates are crystalline complexes of nonpolar compounds with water (usually formed at low temperatures and high pressures) in which the nonpolar molecules are enclosed, as shown, by a polyhedral cage of tetrahedrally hydrogen bonded water molecules (here represented by only their oxygen atoms). The hydrogen bonding interactions of one such water molecule (*arrow*) are shown in detail. [Figure copyrighted © by Irving Geis.]

Walter Kauzmann pointed out in the 1950s that *hydrophobic forces are a major influence in causing proteins to fold into their native conformations.* Figure 7-54 indicates that the amino acid side chain **hydropathies** (indexes of combined hydrophobic and hydrophilic tendencies; Table 7-5) are, in fact, good predictors of which portions of a polypeptide chain are inside a protein, out of contact with

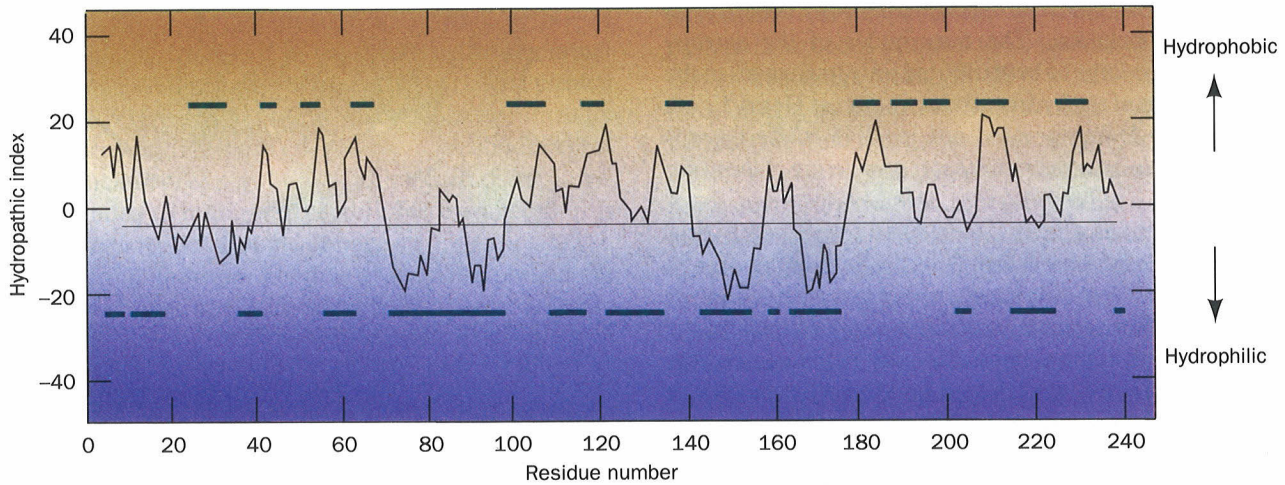


FIGURE 7-54. The hydropathic index (sum of the hydropathies of nine consecutive residues; see Table 7-5) versus the residue sequence number for bovine **chymotrypsinogen.** A large positive hydropathic index is indicative of a hydrophobic region of the polypeptide chain, whereas a large negative value is indicative of a hydrophilic region. The bars above the midpoint line denote the protein's interior regions, as determined by X-ray crystallography, and the bars below the midpoint line indicate the protein's exterior regions. [After Kyte, J. and Doolittle, R.F., *J. Mol. Biol.* **157,** 111 (1982).]

TABLE 7-5. HYDROPATHY SCALE FOR AMINO ACID SIDE CHAINS

Side Chain	Hydropathy
Ile	4.5
Val	4.2
Leu	3.8
Phe	2.8
Cys	2.5
Met	1.9
Ala	1.8
Gly	−0.4
Thr	−0.7
Ser	−0.8
Trp	−0.9
Tyr	−1.3
Pro	−1.6
His	−3.2
Glu	−3.5
Gln	−3.5
Asp	−3.5
Asn	−3.5
Lys	−3.9
Arg	−4.5

Source: Kyte, J. and Doolitle, R.F., *J. Mol. Biol.* **157,** 110 (1982).

the aqueous solvent, and which portions are outside, in contact with the aqueous solvent. In proteins, the effects of hydrophobic forces are often termed **hydrophobic bonding,** presumably to indicate the specific nature of protein folding under the influence of the hydrophobic effect. You should keep in mind, however, that hydrophobic bonding does not generate the directionally specific interactions usually associated with the term "bond."

D. Disulfide Bonds

Since disulfide bonds form as a protein folds to its native conformation (Section 8-1B), they function to stabilize its three-dimensional structure. The relatively reducing chemical character of the cytoplasm, however, greatly diminishes the stability of intracellular disulfide bonds. In fact, almost all proteins with disulfide bonds are secreted to more oxidized extracellular destinations where their disulfide bonds are effective in stabilizing protein structures [secreted proteins fold to their native conformations—and hence form their disulfide bonds—in the endoplasmic reticulum (Section 11-4B) which, unlike other cell compartments, has an oxidizing environment]. Apparently, the relative "hostility" of extracellular environments towards proteins (e.g., uncontrolled temperatures and pH's) requires the additional structural stability conferred by disulfide bonds.

E. Protein Denaturation

The low conformational stabilities of native proteins make them easily susceptible to denaturation by altering the balance of the weak nonbonding forces that maintain the native conformation. When a protein in solution is heated, its conformationally sensitive properties, such as optical rotation (Section 4-2A), viscosity, and UV absorption, change abruptly over a narrow temperature range (Fig. 7-55). *Such a nearly discontinuous change indicates that the native protein structure unfolds in a cooperative manner: Any partial unfolding of the structure destabilizes the remaining structure, which must simultaneously collapse to the random coil.* The temperature at the midpoint of this process is known as the protein's **melting temperature,** T_m, in analogy with the melting of a solid. Most proteins have T_m values well below 100°C. Among the exceptions to this generalization, however, are the proteins of **thermophilic bacteria,** organisms that inhabit hot springs with temperatures approaching 100°C. Interestingly, the X-ray structures of these heat-stable proteins are but subtly different from those of their normally stable homologs.

In addition to high temperatures, proteins are denatured by a variety of other conditions and substances:

1. pH variations alter the ionization states of amino acid side chains (Table 4-1), which changes protein charge distributions and hydrogen bonding requirements.
2. Detergents, some of which significantly perturb protein structures at concentrations as low as $10^{-6}M$, hydrophobically associate with the nonpolar residues of a protein, thereby interfering with the hydrophobic interactions responsible for the protein's native structure.

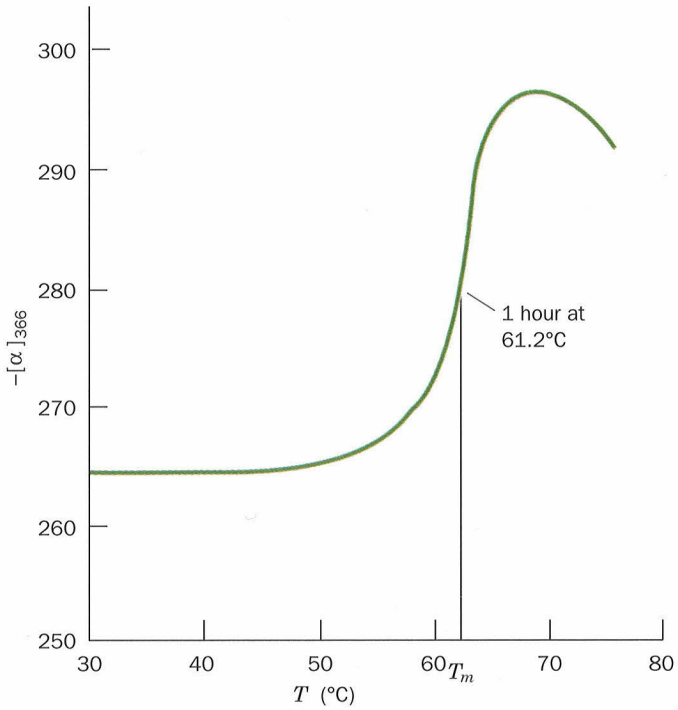
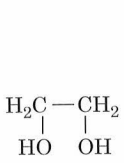
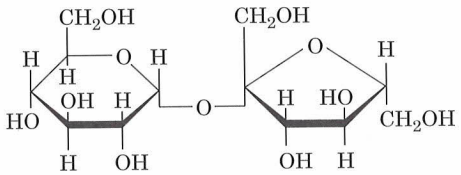


FIGURE 7-55. The optical rotation, at 366 nm, as a function of temperature, of bovine pancreatic ribonuclease A (RNase A) in 0.15M KCl and 0.013M sodium cacodylate buffer, pH 7. The melting temperature, T_m, is defined as the midpoint of the transition. [After von Hippel, P.H. and Wong, K.Y., *J. Biol. Chem.* **10,** 3911 (1965).]

3. High concentrations of water-soluble organic substances, such as aliphatic alcohols, interfere with the hydrophobic forces stabilizing protein structures through their own hydrophobic interactions with water. Organic substances with several hydroxyl groups, such as ethylene glycol or sucrose,



Ethylene glycol



Sucrose

however, are relatively poor denaturants because their hydrogen bonding ability renders them less disruptive of water structure.

The influence of salts is more variable. Figure 7-56 shows the effects of a number of salts on the T_m of bovine pancreatic **ribonuclease A (RNase A).** Some salts, such as $(NH_4)_2SO_4$ and KH_2PO_4, stabilize the native protein structure (raise its T_m); others, such as KCl and NaCl, have little effect; and yet others, such as KSCN and LiBr, destabilize it. The order of effectiveness of the various ions in stabilizing a protein, which is largely independent of the identity of the protein, parallels their capacity to salt out proteins (Section 5-2A). This order is known as the **Hofmeister series:**

Anions $SO_4^{2-} > H_2PO_4^- > CH_3COO^- > Cl^- > Br^- > I^- > ClO_4^- > SCN^-$

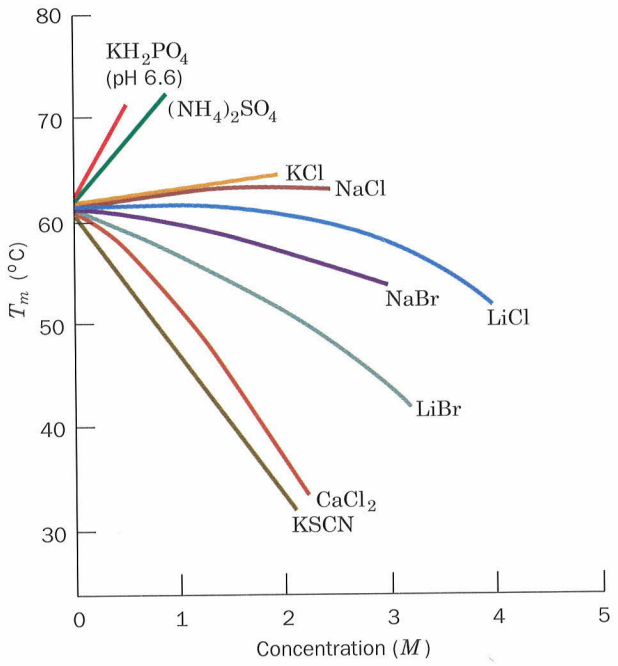
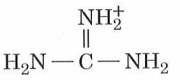


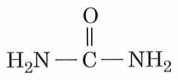
FIGURE 7-56. The melting temperature of RNase A as a function of the concentrations of various salts. All solutions also contain 0.15*M* KCl and 0.013*M* sodium cacodylate buffer, pH 7. [After von Hippel, P.J. and Wong, K.Y., *J. Biol. Chem.* **10,** 3913 (1965).]

Cations $NH_4^+, Cs^+, K^+, Na^+ > Li^+ > Mg^{2+} > Ca^{2+} > Ba^{2+}$

The ions in the Hofmeister series that tend to denature proteins, I^-, ClO_4^-, SCN^-, Li^+, Mg^{2+}, Ca^{2+}, and Ba^{2+}, are said to be **chaotropic.** This list should also include the guanidinium ion (Gu^+) and the nonionic urea,



Guanidinium ion



Urea

which, in concentrations in the range 5 to 10*M*, are the most commonly used protein denaturants. The effect of the various ions on proteins is largely cumulative: GuSCN is a much more potent denaturant than the often used GuCl, whereas Gu_2SO_4 stabilizes protein structures.

Chaotropic agents increase the solubility of nonpolar substances in water. Consequently, their effectiveness as denaturing agents stems from their ability to disrupt hydrophobic interactions although the manner in which they do so is not well understood. Conversely, those substances listed that stabilize proteins strengthen hydrophobic forces, thus increasing the tendency of water to expel proteins. This accounts for the correlation between the abilities of an ion to stabilize proteins and to salt them out.

5. QUATERNARY STRUCTURE

Proteins, because of their multiple polar and nonpolar groups, stick to almost anything; anything, that is, but other proteins. This is because the forces of evolution have arranged the surface groups of proteins so as to prevent their association under physiological conditions. If this were not the case, their resulting nonspecific aggregation would render proteins functionally useless (recall, e.g., the consequences of sickle-cell anemia; Section 6-3A). In his pioneering ultracentrifugational studies on proteins, however, The Svedberg discovered that some proteins are composed of more than one polypeptide chain. Subsequent studies established that this is, in fact, true of most proteins, including nearly all those with molecular masses >100 kD. Furthermore, these polypeptide **subunits** associate in a geometrically specific manner. The spatial arrangement of these subunits is known as a protein's **quaternary structure (4° structure).**

There are several reasons that multisubunit proteins are so common. In large assemblies of proteins, such as collagen fibrils, the advantages of subunit construction over the synthesis of one huge polypeptide chain are analogous to those of using prefabricated components in constructing a building. Defects can be repaired by simply replacing the flawed subunit, the site of subunit manufacture can be different from the site of assembly into the final product, and the only genetic information necessary to specify the entire edifice is that specifying its few different self-assembling subunits. In the case of enzymes, increasing a protein's size

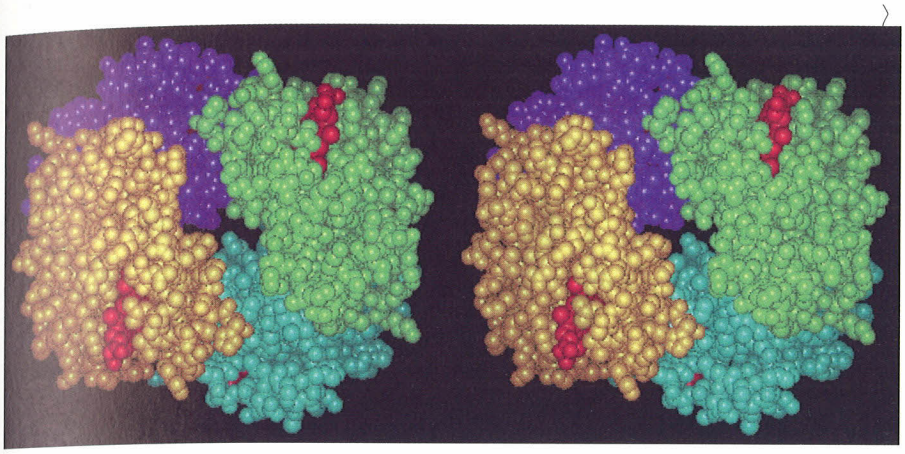


FIGURE 7-57. A stereo, space-filling drawing showing the quaternary structure of hemoglobin. The α_1, α_2, β_1, and β_2 subunits are colored yellow, green, light blue, and purple, respectively. Heme groups are red. The protein is viewed along its molecular twofold rotation axis which relates the $\alpha_1\beta_1$ protomer to the $\alpha_2\beta_2$ protomer. Instructions for viewing stereo drawings are given in the appendix to this chapter.

tends to better fix the three-dimensional positions of the groups forming the enzyme's active site. Increasing the size of an enzyme through the association of identical subunits is more efficient, in this regard, than increasing the length of its polypeptide chain since each subunit has an active site. More importantly, however, the subunit construction of many enzymes provides the structural basis for the regulation of their activities. Mechanisms for this indispensable function are discussed in Sections 9-4 and 12-4.

In this section we discuss how the subunits of multisubunit proteins associate, what sorts of symmetries they have, and how their stoichiometries may be determined.

A. Subunit Interactions

A multisubunit protein may consist of identical or nonidentical polypeptide chains. Recall that hemoglobin, for example, has the subunit composition $\alpha_2\beta_2$. We shall refer to proteins with identical subunits as **oligomers** and to these identical subunits as **protomers.** A protomer may therefore consist of one polypeptide chain or several unlike polypeptide chains. In this sense, hemoglobin is a **dimer** (oligomer of two protomers) of $\alpha\beta$ protomers (Fig. 7-57).

The contact regions between subunits closely resemble the interior of a single subunit protein. They contain closely packed nonpolar side chains, hydrogen bonds involving the polypeptide backbones and their side chains, and, in some cases, interchain disulfide bonds.

B. Symmetry in Proteins

In the vast majority of oligomeric proteins, the protomers are symmetrically arranged; that is, the protomers occupy geometrically equivalent positions in the oligomer. This implies that each protomer has exhausted its capacity to bind to other protomers; otherwise, higher oligomers would form. As a result of this limited binding capacity, protomers pack about a single point to form a closed shell. Proteins cannot have inversion or mirror symmetry, however, because such symmetry operations convert chiral L-residues to D-residues. Thus, *proteins can only have rotational symmetry.*

Various types of rotational symmetry occur in proteins:

1. Cyclic symmetry

In the simplest type of rotational symmetry, **cyclic symmetry,** subunits are related (brought to coincidence) by a single axis of rotation (Fig. 7-58*a*). Objects with 2, 3, . . . , or n-fold rotational axes are said to have C_2, C_3, . . . , or C_n symmetry, respectively. An oligomer

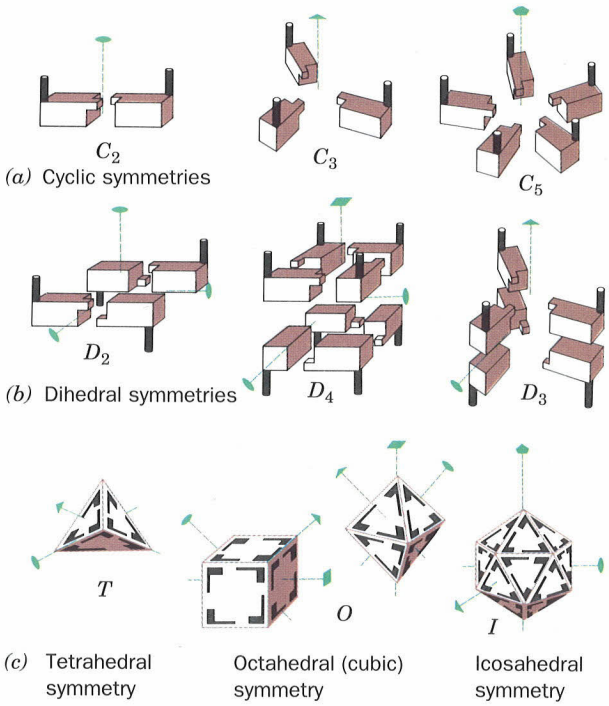


FIGURE 7-58. Some possible symmetries of proteins with identical protomers. The lenticular shape, the triangle, the square, and the pentagon at the ends of the dashed lines indicate, respectively, the unique twofold, threefold, fourfold, and fivefold rotational axes of the objects shown. (*a*) Assemblies with the cyclic symmetries C_2, C_3, and C_5. (*b*) Assemblies with the dihedral symmetries D_2, D_4, and D_3. In these objects, a twofold axis is perpendicular to the vertical two-, four-, and threefold axes. (*c*) Assemblies with T, O, and I symmetry. Note that the tetrahedron has some but not all of the symmetry elements of the cube, and that the cube and the octahedron have the same symmetry. [Figure copyrighted © by Irving Geis.]

with C_n symmetry consists of n protomers that are related by $(360/n)°$ rotations. C_2 symmetry is the most common symmetry in proteins; higher cyclic symmetries are relatively rare.

A common mode of association between protomers related by a twofold rotation axis is the continuation of a β sheet across subunit boundaries. In such cases, the twofold axis is perpendicular to the β sheet so that two symmetry equivalent strands hydrogen bond in an antiparallel fashion. In this manner, the sandwich of two four-stranded β sheets of the **prealbumin** protomer is extended across a twofold axis to form a sandwich of two eight-stranded β sheets (Fig. 7-59).

2. **Dihedral symmetry**

Dihedral symmetry (D_n), a more complicated type of rotational symmetry, is generated when an n-fold rotation axis and a twofold rotation axis intersect at right angles (Fig. 7-58*b*). An oligomer with D_n symmetry consists of $2n$ protomers. The D_2 symmetry is, by far, the most common type of dihedral symmetry in proteins. Under the proper conditions, many oligomers with D_n symmetry will dissociate into two oligomers, each with C_n symmetry (and which were related by the twofold rotation axis in the D_n oligomer). These, in turn, dissociate to their component protomers under more stringent dissociating conditions.

3. **Other rotational symmetries**

The only other types of rotationally symmetric objects are those that have the rotational symmetries of a tetrahedron (T), a cube or octahedron (O), or an icosahedron (I), and have 12, 24, and 60 equivalent positions, respectively (Fig. 7-58*c*). The subunit arrangements in the protein coats of the so-called spherical viruses are based on icosahedral symmetry (Section 32-2A).

Under favorable conditions electron microscopy can provide dramatic indications of oligomeric symmetry.

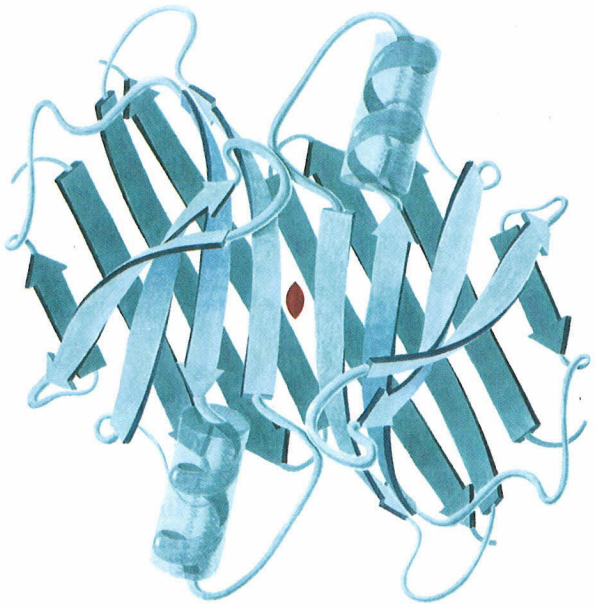


FIGURE 7-59. A prealbumin dimer viewed down its twofold axis (*red symbol*). Each protomer consists of a sandwich of two four-stranded β sheets. Note how both of these β sheets are continued in an antiparallel fashion in the other protomer to form a sandwich of two eight-stranded β sheets. Two of these dimers associate back to back in the native protein to form a tetramer with D_2 symmetry. [After a drawing by Jane Richardson, Duke University.]

Electron microscopy studies suggest, for example, that the 600 kD *E. coli* **glutamine synthetase** has D_6 symmetry (Fig. 7-60). Unfortunately, since this technique has insufficient resolution to reveal the relative orientations of the protein subunits (i.e., the directions of the arrows in the interpretive drawing of Fig. 7-60), such symmetry assignments must be taken as tentative; only X-ray crystal structure analysis can unambiguously establish the geometric relationships

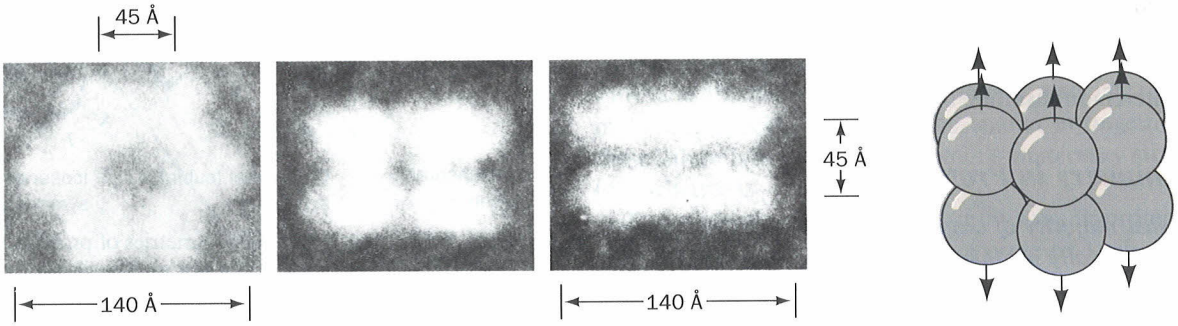


FIGURE 7-60. Sets of five superimposed electron micrographs (to enhance real detail) of *E. coli* glutamine synthetase molecules in their three characteristic orientations. The mean dimensions are indicated. When the oligomeric molecule rests on its face, it appears to be a hexagonal ring of subunits (*left*). Molecules on edge, however, show two layers of subunits as four spots when viewed exactly between the subunits (*middle*), or as two parallel streaks when viewed in other directions (*right*). This suggests, as the accompanying drawing indicates, that the enzyme molecule has 12 identical subunits organized with D_6 symmetry into two hexagons that are stacked with their subunits in apposition. [Courtesy of Earl Stadtman, NIH.]

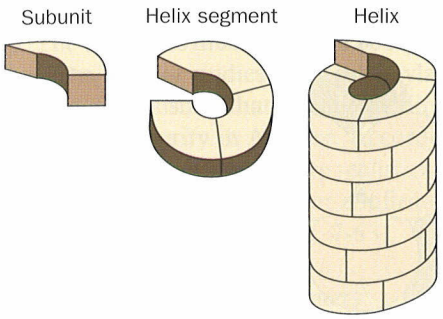


FIGURE 7-61. A helical structure composed of a single kind of subunit.

among protein subunits. In the case of glutamine synthetase, however, X-ray studies have confirmed that it indeed has D_6 symmetry (Section 24-5A).

Helical Symmetry

Some protein oligomers have **helical symmetry** (Fig. 7-61). The chemically identical subunits in a helix are not strictly equivalent because, for instance, those at the end of the helix have a different environment than those in the middle. Nevertheless, the surroundings of all subunits in a long helix, except those near its ends, are sufficiently similar that the subunits are said to be **quasi-equivalent.** The subunits of many structural proteins, for example, those of muscle (Section 34-3A), assemble into fibers with helical symmetry.

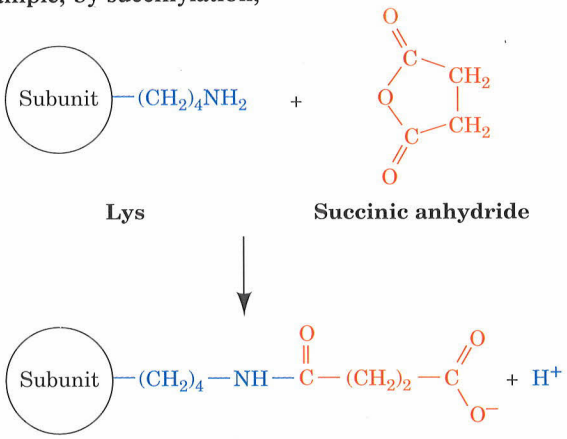
C. Determination of Subunit Composition

The number of different types of subunits in an oligomeric protein may be determined by end group analysis (Section 6-1A). In principle, the subunit composition of a protein may be determined by comparing its molecular mass with those of its component subunits. In practice, however, experimental difficulties, such as the partial dissociation of a supposedly intact protein and uncertainties in molecular mass determinations, often provide erroneous results.

Hydridization Yields Quaternary Structural Information

An alternative procedure may be used if two chemically different and therefore separable species of the protein are available. The species may be proteins with slightly different 1° structures from different organisms or, as is often the case, variants of a protein that occur in the same organism. The two different oligomeric proteins are purified, mixed together, dissociated to their component subunits by exposure to mildly denaturing conditions (e.g., by changing the pH or adding urea), and then allowed to reassemble (e.g., by restoring the pH or dialyzing out the urea). If the native proteins are n-mers, S_n and S'_n, this procedure will yield $(n + 1)$ species of **hybrid molecules** with the mixed subunit compositions S_n, $S_{n-1}S'$, $S_{n-2}S'_2$, . . . , S'_n, which can be analyzed, for example, by electrophoresis. For instance, vertebrates possess two varieties of the enzyme **lactate dehydrogenase (LDH):** the M type, which predominates in skeletal muscle, and the H type, which predominates in heart tissue. Hybridization of these oligomers, in this case by repeated freezing and thawing, yields five **isozymes** (isoenzymes; catalytically and structurally similar enzymes from the same organism) of LDH that have the subunit compositions M_4, M_3H, M_2H_2, MH_3, and H_4 (Fig. 7-62). This demonstrates that LDH is a tetramer.

In a related method, a protein subunit may be labeled, for example, by succinylation,



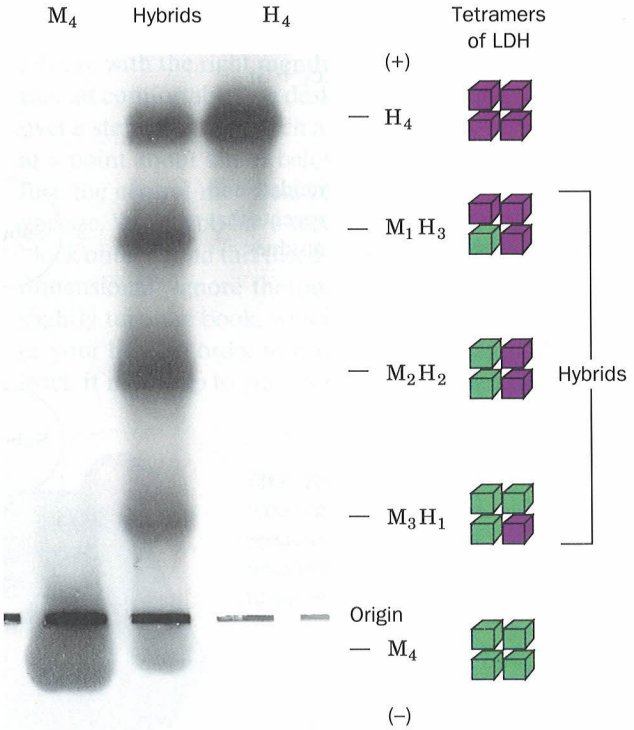
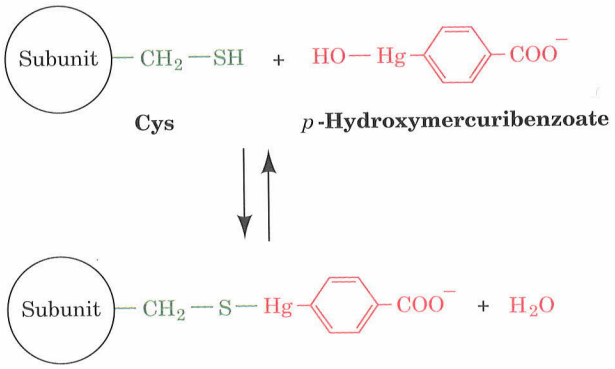


FIGURE 7-62. An electrophoretogram of bovine lactate dehydrogenase. The M and H forms of LDH (*outer lanes*) have different electrophoretic mobilities. Upon hybridization of these oligomers, five electrophoretically distinct isozymes are formed (*center lane*), which indicates that LDH is a tetramer. [Courtesy of Clement Markert, North Carolina State University at Raleigh.]

which alters the electrophoretic mobility of a protein by changing its ionic charge. John Gerhart and Howard Schachman used this technique to determine the geometric distribution of subunits in *E. coli* **aspartate transcarbamoylase (ATCase).** ATCase has two types of subunits, the catalytic subunit, c, and the regulatory subunit, r (their enzymatic roles are discussed in Section 12-4). Molecular mass measurements (c = 33 kD, r = 17 kD, and ATCase = 300 kD) indicate that ATCase has the subunit composition c_6r_6. This was corroborated by preliminary X-ray studies which established that ATCase has D_3 symmetry (recall that a protein of D_3 symmetry must have six protomers; Fig. 7-58*b*). Treatment with organic mercurials such as ***para*-hydroxymercuribenzoate,** which reacts with Cys sulfhydryl groups



causes ATCase to dissociate according to the reaction

$$c_6r_6 \rightarrow 2c_3 + 3r_2$$

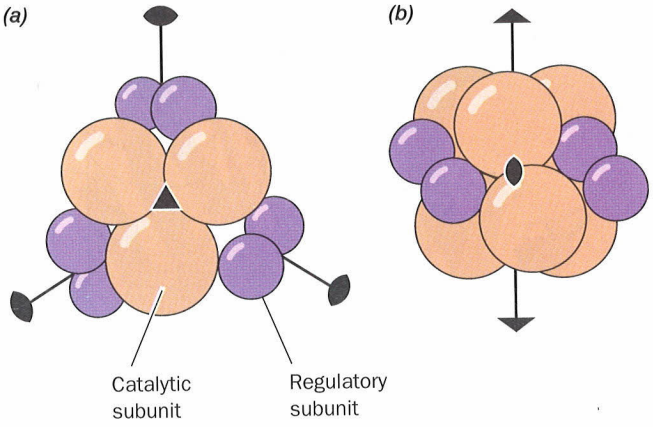
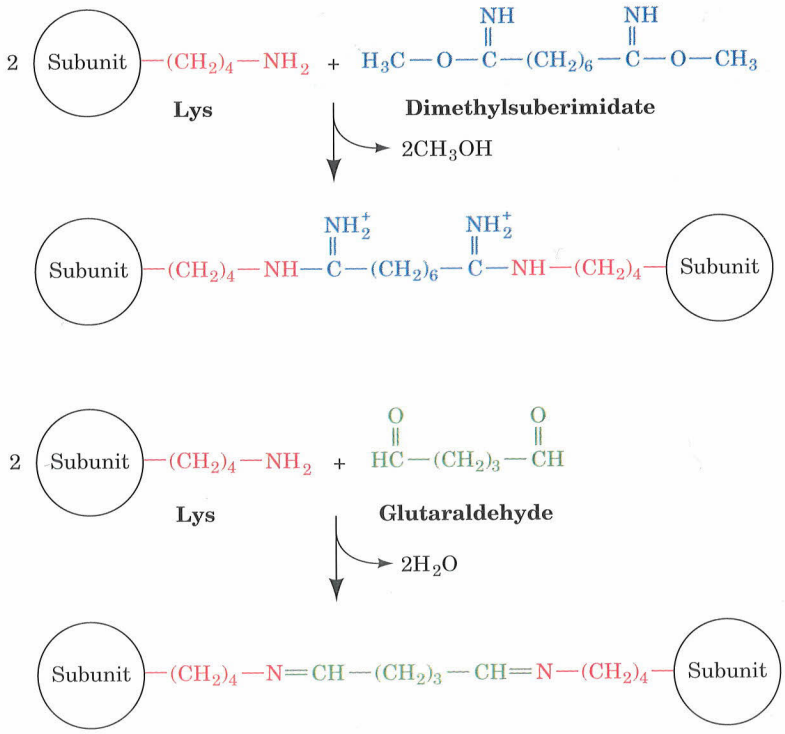


FIGURE 7-63. The quaternary structure of *E. coli* aspartate transcarbamoylase as established by X-ray structure analysis. Catalytic subunits and regulatory subunits are represented, respectively, by large orange spheres and small purple spheres. The molecule has D_3 symmetry. (*a*) View along the threefold axis (*triangle*). (*b*) View along a twofold axis (*lenticular shapes*). [After Kantrowitz, E. R., Pastra-Landis, S.C., and Lipscomb, W.N., *Trends Biochem. Sci.* **5,** 150 (1980).]

The catalytic trimers, c_3, were isolated and succinylated to form c_3^s. When these were mixed with unmodified catalytic trimers and excess regulatory dimers, r_2, under conditions that ATCase reforms, only three products could be electrophoretically distinguished: c_6r_6, $c_3c_3^sr_6$, and $c_6^sr_6$. This indicates that the catalytic subunits were not exchanged be-

FIGURE 7-64. Dimethylsuberimidate and glutaraldehyde are bifunctional reagents that react to covalently cross-link two Lys residues.



tween catalytic trimers in ATCase; for example, no $c_4c_2^sr_6$ was formed. The c_3 trimers must therefore be separate entities in the enzyme. Similar studies using succinylated regulatory dimers, r_2^s, established that regulatory dimers likewise maintain their integrity in ATCase. Accordingly, the subunit composition of ATCase is more realistically represented as $(c_3)_2(r_2)_3$. This result was later confirmed by the X-ray crystal structure of ATCase (Fig. 7-63).

Cross-Linking Agents Stabilize Oligomers

Another method for 4° structure analysis, which is especially useful for oligomeric proteins that decompose easily, employs **cross-linking agents** such as **dimethylsuberimidate** or **glutaraldehyde** (Fig. 7-64). If carried out at sufficiently low protein concentrations to eliminate intermolecular reactions, cross-linking reactions will covalently join only the subunits in a molecule that are no further apart than the length of the cross-link (assuming, of course, that the proper amino acid residues are present). The molecular mass of a cross-linked protein therefore places a lower limit on its number of subunits. Such studies can also provide some indication of the distance between subunits, particularly if a series of cross-linking agents with different lengths is employed.

APPENDIX: VIEWING STEREO PICTURES

Although we live in a three-dimensional world, the images that we see have been projected onto the two-dimensional plane of our retinas. Depth perception therefore involves binocular vision: The slightly different views perceived by each eye are synthesized by the brain into a single three-dimensional impression.

Two-dimensional pictures of complex three-dimensional objects are difficult to interpret because most of the information concerning the third dimension is suppressed. This information can be recovered by presenting each eye with the image only it would see if the three-dimensional object were actually being viewed. A **stereo pair** therefore consists of two images, one for each eye. Corresponding points of stereo pairs are generally separated by ~6 cm, the average distance between human eyes. Stereo drawings are usually computer generated because of the required precision of the geometric relationship between the members of a stereo pair.

In viewing a stereo picture, one must overcome the visual habits of a lifetime because each eye must see its corresponding view independently. Viewers are commercially available to aid in this endeavor. However, with some training and practice, equivalent results can be obtained without their use.

To train yourself to view stereo pictures, you should become aware that each eye sees a separate image. Hold your finger up about a foot (30 cm) before your eyes while fixing your gaze on some object beyond it. You may realize that you are seeing two images of your finger. If, after some concentration, you are still aware of only one image, try blinking your eyes alternately to ascertain which of your eyes is seeing the image you perceive. Perhaps alternately covering and uncovering this dominant eye while staring past your finger will help you become aware of the independent workings of your eyes.

The principle involved in seeing a stereo picture is to visually fuse the left member of the stereo pair seen by the left eye with the right member seen by the right eye. To do this, sit comfortably at a desk, center your eyes about a foot over a stereo drawing such as Fig. 7-65 and stare through it at a point about a foot below the drawing. Try to visually fuse the central members of the four out-of-focus images you see. When you have succeeded, your visual system will "lock onto" it and this fused central image will appear three dimensional. Ignore the outer images. You may have to slightly turn the book, which should be held perfectly flat, or your head in order to bring the two images to the same level. It may help to place the book near the edge of a desk,

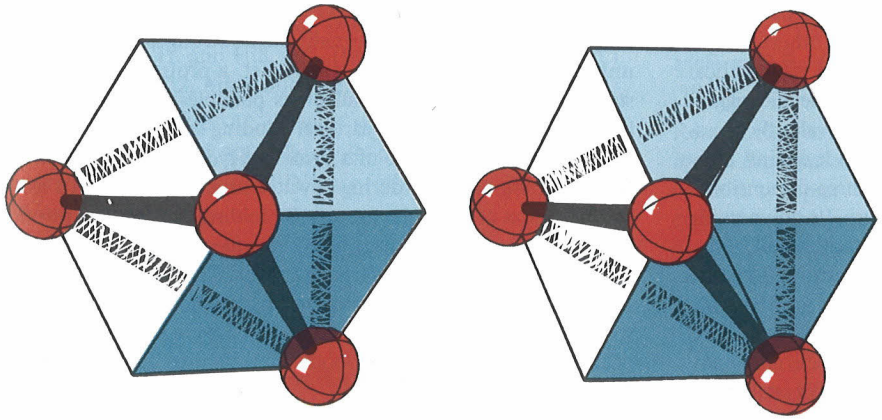


FIGURE 7-65. A stereo drawing of a tetrahedron inscribed in a cube. When properly viewed, the apex of the tetrahedron should appear to be pointing towards the viewer.

center your finger about a foot below the drawing, and fixate on your finger while concentrating on the stereo pair. Another trick is to hold your flattened hand or an index card between your eyes so that the left eye sees only the left half of the stereo pair and the right eye sees only the right half and then fuse the two images you see.

The final step in viewing a stereo picture is to focus on the image while maintaining fusion. This may not be easy because our ingrained tendency is to focus on the point at which our gaze converges. It may help to move your head closer to or further from the picture. Most people (including the authors) require a fair amount of practice to become proficient at seeing stereo without a viewer. However, the three-dimensional information provided by stereo pictures, not to mention their esthetic appeal, makes it worth the effort. In any case, the few stereo figures used in this text have been selected for their visual clarity without the use of stereo; stereo will simply enhance their impression of depth.

CHAPTER SUMMARY

The peptide group is constrained by resonance effects to a planar, trans conformation. Steric interactions further limit the conformations of the polypeptide backbone by restricting the torsion angles, ϕ and ψ, of each peptide group to three small regions of the Ramachandran diagram. The α helix, whose conformation angles fall within the allowed regions of the Ramachandram diagram, is held together by hydrogen bonds. The 3_{10} helix, which is more tightly coiled than the α helix, lies in a mildly forbidden region of the Ramachandran diagram. Its infrequent occurrences are most often as single-turn terminations of α helices. In the parallel and antiparallel β pleated sheets, two or more almost fully extended polypeptide chains associate such that neighboring chains are hydrogen bonded. These β sheets have a right-handed curl when viewed along their polypeptide chains. The polypeptide chain often reverses its direction through a β bend. Other arrangements of the polypeptide chain, which are collectively known as coil conformations, are more difficult to describe but are no less ordered than are α or β structures.

The mechanical properties of fibrous proteins can often be correlated with their structures. Keratin, the principal component of hair, horn, and nails, forms protofibrils that consist of two pairs of α helices in which the members of each pair are twisted together into a left-handed coil. The pliability of keratin decreases as the content of disulfide cross-links between the protofibrils increases. Silk fibroin forms flexible but inextensible fibers of great strength. It exists as a semicrystalline array of antiparallel β sheets in which layers of Gly side chains alternate with layers of Ala and Ser side chains. Collagen is the major protein component of connective tissue. Its every third residue is Gly and many of the others are Pro and Hyp. This permits collagen to form a ropelike triple helical structure that has great tensile strength. Collagen molecules aggregate in a staggered array to form fibrils that are covalently cross-linked by groups derived from their His and Lys side chains. Elastin, which has elastic properties, forms a three-dimensional network of fibers that exhibit no regular structure. Its polypeptide strands are cross-linked in a manner similar to that in collagen.

The accuracies of protein X-ray structure determinations are limited by crystalline disorder to resolutions that are mostly in the range 2.0 to 3.5 Å. This requires that a protein's structure be determined by fitting its primary structure to its electron density map. Several lines of evidence indicate that protein crystal structures are nearly identical to their solution structures. The structures of small proteins may also be determined in solution by 2D-NMR techniques which, for the most part, yield results similar to those of X-ray crystal structures. A globular protein's 3° structure is the arrangement of its various elements of 2° structure together with the spatial dispositions of its side chains. Its amino acid residues tend to segregate according to residue polarity. Nonpolar residues preferentially occur in the interior of a protein out of contact with the aqueous solvent, whereas charged polar residues are located on its surface. Uncharged polar residues may occur at either location but, if they are internal, they form hydrogen bonds with other protein groups. The interior of a protein molecule resembles a crystal of an organic molecule in its packing efficiency. Larger proteins often fold into two or more domains that may have functionally and structurally independent properties. Certain groupings of secondary structural elements, known as supersecondary structures, repeatedly occur as components of globular proteins. They may have functional as well as structural significance.

Proteins have marginally stable native structures that form as a result of a fine balance among the various noncovalent forces to which they are subject: ionic and dipolar interactions, hydrogen bonding, and hydrophobic forces. Ionic interactions are relatively weak in aqueous solutions due to the solvating effects of water. The various interactions among permanent and induced dipoles, which are collectively referred to as van der Waals forces, are even weaker and are effective only at short range. Nevertheless, because of their large numbers, they cumulatively have an important influence on protein structures. Hydrogen bonding forces are far more directional than are other noncovalent forces. They add little stability to a protein structure, however, because the hydrogen bonds that native proteins form internally are no stronger than those that unfolded proteins form with water. Yet, a protein can only fold stably in ways that almost all of its possible internal hydrogen bonds are formed so that hydrogen bonding is important in specifying the native structure of a protein. Hydrophobic forces arise from the unfavorable ordering of water structure that results from the hydration of nonpolar groups. By folding such that its nonpolar groups are out of contact with the aqueous solvent, a protein minimizes these unfavorable interactions. The fact that most protein denaturants interfere with the hydrophobic effect demonstrates the importance of hydrophobic forces in stabilizing native protein structures. Disulfide bonds often stabilize the native structures of extracellular proteins.

Many proteins consist of noncovalently linked aggregates of subunits in which the subunits may or may not be identical. Most

oligomeric proteins are rotationally symmetric. The protomers in many fibrous proteins are related by helical symmetry. The subunit structures of proteins may be elucidated by a variety of techniques, including the hybridization of subunits with their naturally occurring or derivatized variants, cross-linking studies, electron microscopy, and X-ray crystal structure analysis.

REFERENCES

General

Branden, C. and Tooze, J., *Introduction to Protein Structure,* Garland Publishing (1991).

Cantor, C.R. and Schimmel, P.R., *Biophysical Chemistry,* Chapters 2 and 5, Freeman (1980).

Creighton, T.E., *Proteins* (2nd ed.), Chapters 4–6, Freeman (1993).

Dickerson, R.E. and Geis, I., *The Structure and Action of Proteins,* Benjamin/Cummings (1969). [A marvelously illustrated exposition of the fundamentals of protein structure.]

Perutz, M., *Protein Structure. New Approaches to Disease and Therapy,* Freeman (1992). [A series of short articles on the structures of a variety of proteins and their biomedical implications.]

Schultz, G.E. and Schirmer, R.H., *Principles of Protein Structure,* Chapters 2–5 and 7, Springer–Verlag (1979). [An advanced text.]

Wood, W.B., Wilson, J.H., Benbow, R.M., and Hood, L.E., *Biochemistry. A Problems Approach* (2nd ed.), Chapters 4 and 5, Benjamin/Cummings (1981). [A question-and-answer approach to learning protein structures.]

Secondary Structure

Leszczynski, J.F. and Rose, G.D., Loops in globular proteins: a novel category of secondary structure, *Science* **234,** 849–855 (1986).

Toniolo, C. and Benedetti, E.,The polypeptide 3_{10}-helix, *Trends Biochem. Sci.* **16,** 350–353 (1991).

Rose, G.D., Gierasch, L.M., and Smith, J.A., Turns in polypeptides and proteins, *Adv. Protein Chem.* **37,** 1–109 (1985).

Salemme, F.R., Structural properties of protein β-sheets, *Prog. Biophys. Mol. Biol.* **42,** 95–133 (1983).

Fibrous Proteins

Bornstein, P. and Traub, W., The chemistry and biology of collagen, *in* Neurath, H. and Hill, R.L. (Eds.), *The Proteins* (3rd ed.), Vol. 4, *pp.* 412–632, Academic Press (1979).

Byers, P.H., Disorders of collagen synthesis and structure, *in* Scriver, C.R., Beaudet, A.L., Sly, W.S., and Valle, D. (Eds.), *The Metabolic Basis of Inherited Disease* (6th ed.), *pp.* 2805–2842, McGraw–Hill (1989).

Engel, J. and Prokop, D.J., The zipper-like folding of collagen triple helices and the effects of mutations that disrupt the zipper, *Annu. Rev. Biophys. Biophys. Chem.* **20,** 137–152 (1991).

Eyre, D.R., Paz, M.A., and Gallop, P.M., Cross-linking in collagen and elastin, *Annu. Rev. Biochem.* **53,** 717–748 (1984).

Fuchs, E. and Coulombe, P.A., Of mice and men: Genetic skin diseases of keratin, *Cell* **69,** 899–902 (1992).

Jones, E.Y. and Miller, A., Analysis of structural design features in collagen, *J. Mol. Biol.* **218,** 209–219 (1991).

Kaplan, D., Adams, W.W., Farmer, B., and Viney, C., *Silk Polymers,* Am. Chem. Soc. (1994).

Martin, G.R., Timpl, R., Müller, P.K., and Kühn, K., The genetically distinct collagens, *Trends Biochem. Sci.* **10,** 285–287 (1985).

Nimni, M.E. (Ed.), *Collagen,* Vol. I, CRC Press (1988). [Contains articles on collagen biochemistry.]

Prockop, D.J., Mutations in collagen genes as a cause of connective-tissue disease, *New Engl. J. Med.* **326,** 540–546 (1992).

Robert, L. and Hornebeck, W. (Eds.), *Elastin and Elastases,* Vol. 1, CRC Press (1989).

Steinert, P.M. and Parry, D.A.D., Intermediate filaments, *Annu. Rev. Cell Biol.* **1,** 41–65 (1985). [Discusses the structure of a keratin.]

van der Rest, M. and Bruckner, P., Collagens: diversity at the molecular and supramolecular levels, *Curr. Opin. Struct. Biol.* **3,** 430–436 (1993).

Vuorio, E. and de Crombrugghe, B., The family of collagen genes, *Annu. Rev. Biochem.* **59,** 837–872 (1990). [Discusses the various types of collagens.]

Globular Proteins

Chothia, C., Principles that determine the structures of proteins, *Annu. Rev. Biochem.* **53,** 537–572 (1984).

Clore, G.M. and Gronenborn, A.M., Two-, three-, and four-dimensional NMR methods for obtaining larger and more precise three-dimensional structures of proteins in solution, *Annu. Rev. Biophys. Biophys. Chem.* **20,** 29–63 (1991).

Cohen, C. and Parry, D.A.D., α-Helical coiled coils and bundles: How to design an α-helical protein, *Proteins* **7,** 1–15 (1990).

Glusker, J.P., Lewis, M., and Rossi, M., *Crystal Structure Analysis for Chemists and Biologists,* VCH Publishers (1994).

Lesk, A.M., *Protein Architecture. A Practical Approach,* IRL Press (1991).

McRee, D.E., *Practical Protein Cystallography,* Academic Press (1993).

Rees, A.R., Sternberg, M.J.E., and Wetzel, R., *Protein Engineering. A Practical Approach,* IRL Press (1992). [Chapters 1 and 2 contain synopses of protein structure determination by X-ray crystallographic and NMR methods.]

Rhodes, G., *Crystallography Made Clear: A Guide for Users of Macromolecular Models,* Academic Press (1993).

Richards, F.M., Areas, volumes, packing, and protein structure, *Annu. Rev. Biophys. Bioeng.* **6,** 151–176 (1977).

Richardson, J.S. and Richardson, D.C., Principles and patterns of protein conformation, *in* Fasman, G.D. (Ed.), *Prediction of Protein Structure and the Principles of Protein Conformation, pp.* 1–

98, Plenum Press (1989). [A comprehensive account of protein conformations based on X-ray structures.]

Richardson, J.S., The anatomy and taxonomy of protein structures, *Adv. Protein Chem.* **34,** 168–339 (1981). [A detailed discussion of the structural principles governing globular proteins accompanied by an extensive collection of their cartoon representations.]

Thornton, J.M., Protein structures: The end point of the folding pathway, *in* Creighton, T.E. (Ed.), *Protein Folding, pp.* 59–81, Freeman (1992).

Wagner, G., Hyberts, S.G., and Havel, T.F., NMR structure determination in solution. A critique and comparison with X-ray crystallography, *Annu. Rev. Biophys. Biomol. Struct.* **21,** 167–198 (1992).

Wüthrich, K., Protein structure determination in solution by nuclear magnetic resonance spectroscopy, *Science* **243,** 45–50 (1989).

Wyckoff, H.W., Hirs, C.H.W., and Timasheff, S.N. (Eds.), *Diffraction Methods for Biological Macromolecules,* Parts A and B, *Methods Enzymol.* **114, 115** (1985). [A series of articles on the theory and practice of X-ray crystallography.]

Protein Stability

Alber, T., Stabilization energies of protein conformation, *in* Fasman, G.D. (Ed.), *Prediction of Protein Structure and the Principles of Protein Conformation, pp.* 161–192, Plenum Press (1989).

Burley, S.K. and Petsko, G.A., Weakly polar interactions in proteins, *Adv. Protein Chem.* **39,** 125–189 (1988).

Creighton, T.E., Stability of folded proteins, *Curr. Opin. Struct. Biol.* **1,** 5–16 (1991).

Edsall, J.T. and McKenzie, H.A., Water and proteins, *Adv. Biophys.* **16,** 51–183 (1983).

Eigenbrot, C. and Kossiakoff, A.A., Structural consequences of mutation, *Curr. Opin. Biotech.* **3,** 333–337 (1992).

Fersht, A.R. and Serrano, L., Principles of protein stability derived from protein engineering experiments, *Curr. Opin. Struct. Biol.* **3,** 75–83 (1993). [Discusses how the roles of specific side chains in proteins can be quantitatively determined by mutationally changing them and calorimetrically measuring the stabilities of the resulting proteins.]

Harvey, S.C., Treatment of electrostatic effects in macromolecular modeling, *Proteins* **5,** 78–92 (1989).

Jeffrey, G.A. and Saenger, W., *Hydrogen Bonding in Biological Structures,* Springer–Verlag (1991).

Kauzmann, W., Some factors in the interpretation of protein denaturation, *Adv. Protein Chem.* **14,** 1–63 (1958). [A classic review that first pointed out the importance of hydrophobic bonding in stabilizing proteins.]

Matthews, B.W., Structural and genetic analysis of protein stability, *Annu. Rev. Biochem.* **62,** 139–160 (1993).

Ramachandran, G.N. and Sasisekharan, V., Conformation of polypeptides and proteins, *Adv. Protein Chem.* **23,** 283–437 (1968).

Richards, F.M., Folded and unfolded proteins: An introduction, *in* Creighton, T.E. (Ed.), *Protein Folding, pp.* 1–58, Freeman (1992).

Saenger, W., Structure and dynamics of water surrounding biomolecules, *Annu. Rev. Biophys. Biophys. Chem.* **16,** 93–114 (1987).

Schellman, J.A., The thermodynamic stability of proteins, *Annu. Rev. Biophys. Biophys. Chem.* **16,** 115–137 (1987).

Stickle, D.F., Presta, L.G., Dill, K.A., and Rose, G.D., Hydrogen bonding in globular proteins, *J. Mol. Biol.* **226,** 1143–1159 (1992).

Tanford, C., *The Hydrophobic Effect: Formation of Micelles and Biological Membranes* (2nd ed.), Chapters. 2–4 and 13, Wiley (1980).

Teeter, M.M., Water-protein interactions: Theory and experiment, *Annu. Rev. Biophys. Biophys. Chem.* **20,** 577–600 (1991).

Yang, A.-S. and Honig, B., Electrostatic effects on protein stability, *Curr. Opin. Struct. Biol.* **2,** 40–45 (1992).

Quaternary Structure

Eisenstein, E. and Schachman, H.K., Determining the roles of subunits in protein function, *in* Creighton, T.E. (Ed.), *Protein Function. A Practical Approach, pp.* 135–176, IRL Press (1989).

Klotz, I.M., Darnell, D.W., and Langerman, N.R., Quaternary structure of proteins, *in* Neurath, H. and Hill, R.L. (Eds.), *The Proteins* (3rd ed.), Vol. 1, *pp.* 226–411, Academic Press (1975).

Matthews, B.W. and Bernhard, S.A., Structure and symmetry in oligomeric proteins, *Annu. Rev. Biophys. Bioeng.* **6,** 257–317 (1973).

PROBLEMS

1. What is the length of an α helical section of a polypeptide chain of 20 residues? What is its length when it is fully extended (all trans)?

***2.** From an examination of Figs. 7-7 and 7-8, it is apparent that the polypeptide conformation angle ϕ is more constrained than is ψ. By referring to Fig. 7-4, or better yet, by examining a molecular model, indicate the sources of the steric interference that limit the allowed values of ϕ when $\psi = 180°$.

3. For a polypeptide chain made of γ-amino acids, state the nomenclature of the helix analogous to the 3_{10} helix of α-amino acids. Assume the helix has a pitch of 9.9 Å and a rise per residue of 3.2 Å.

***4.** Table 7-6 gives the torsion angles, ϕ and ψ, of hen egg white lysozyme for residues 24–73 of this 129-residue protein. (a) What is the secondary structure of residues 26–35? (b) What is the secondary structure of residues 42–53? (c) What is the